From Dept. Medical Biochemistry and Biophysics
Karolinska Institute, Stockholm, Sweden

IDENTIFYING GENETIC DETERMINANTS OF T CELL-DEPENDENT AUTOIMMUNITY USING FORWARD GENETICS

Gonzalo Fernandez Lahore

Karolinska Institutet

Stockholm, August 2021

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2021

ISBN **978-91-8016-294-4**
Cover illustration: Genetic predisposition determines rheumatoid arthritis. Illustration by the author.

IDENTIFYING GENETIC DETERMINANTS OF T CELL-DEPENDENT AUTOIMMUNITY USING FORWARD GENETICS

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Gonzalo Fernandez Lahore

The thesis will be defended in public at Gustaf Retzius hall, KI, Stockholm, Aug 20, 2021

Principal Supervisor:
Prof. Rikard Holmdahl
Karolinska Institute
Department of Medical Biochemistry and Biophysics
Division of Medical Inflammation Research

Co-supervisor(s):
Prof. Kutty Selva Nandakumar
Southern Medical University
School of Pharmaceutical Sciences

Dr. Liselotte Bäckdahl
Karolinska Institute
Department of Medical Biochemistry and Biophysics
Division of Medical Inflammation Research

Opponent:
Prof. Vijay Kuchroo
Evergrande Center for Immunologic Diseases, Harvard Medical School and Brigham and Women's Hospital

Examination Board:
Assoc. Prof. Lisa Westerberg
Karolinska Institute
Department of Microbiology, Tumor and Cell Biology

Asst. Prof. Jonathan Coquet
Karolinska Institute
Department of Microbiology, Tumor and Cell Biology

Prof. Göran Andersson
Swedish University of Agricultural Sciences
Department of Animal Breeding and Genetics

Como no estas experimentado en las cosas del mundo, todas las cosas que tienen algo de dificultad te parecen imposibles. Confía en el tiempo, que suele dar dulces salidas a muchas amargas dificultades.

Miguel de Cervantes Saavedra —
Don Quijote de la mancha

POPULAR SCIENCE SUMMARY

Rheumatoid arthritis (RA) is a common autoimmune disease of the joints affecting 0,5-1% of the population. People affected by RA suffer from joint pain and have a reduced quality of life. Current treatment options can control RA progression, but do not cure the disease. Thus, RA is a problem for the persons affected by it but also from a socioeconomic standpoint. To develop improved treatment options, it is necessary to gain better understanding of the disease and its causes. Research suggests that susceptibility to arthritis is determined to a large extent by the natural variability in our genes. Understanding RA therefore means identifying and understanding the genes driving it.

Unfortunately, identifying which genes are important for RA and understanding how they work is not a straightforward endeavour. There is a high variability in the human population which complicates data analysis and interpretation. Additionally, there is limited availability of relevant human samples, which makes it difficult to test new ideas and concepts in meaningful experiments. The immune system in mice largely mimics the one in humans and it is possible to induce an RA-like arthritis in rodents. This means that we can study genes regulating arthritis in mice to then transfer our findings to humans.

In this study, we set out to identify RA susceptibility genes using mouse models. We bred arthritis resistant with arthritis susceptible mice, and then evaluated arthritis susceptibility in the offspring. Because the genome of the offspring is a random mosaic of both parental genomes, we can deduce which are the genes that affect arthritis susceptibility by testing enough mice. Using this approach, we herein identify three genetic regions regulating arthritis in mice, and likely also in humans. Our findings contribute to the general understanding of RA genetics and we even identify a new potential therapeutic target.

A summary of our main findings:

1. In the first study, we find that natural genetic variability enhancing the availability of the vitamin D receptor regulates arthritis. Blood vitamin D levels have long been speculated to influence susceptibility to several autoimmune diseases in humans and our findings support the idea that vitamin D is an important regulator of the immune system.

2. In the second study, we identify the gene *SH3GL1* as a novel regulator of autoimmune arthritis. The *SH3GL1* gene serves as the template for the Endophilin A2 protein, which regulates vesicle formation in neural cells. For the first time, we show that this function is also highly relevant to immune cells.

3. In the third study, we find that a genetic region that is important for estrogen-mediated regulation of the T cell marker CD2 regulates arthritis specifically in female mice. These findings could help explain why women are more susceptible to autoimmune diseases, a phenomenon that remains poorly understood.

ABSTRACT

Rheumatoid arthritis (RA) affects 0.5-1% of the population and is an important health and socioeconomic problem. RA has a high degree of heritability. Thus, extensive efforts have been made to better understand the genetic variability contributing to disease susceptibility. However, dissecting the genetic component of RA in humans has been difficult due to heterogeneity in the human population, multiple testing issues, and lack of accessibility to relevant tissues for proof-of-concept studies. Genetic studies in mouse model systems circumvent these problems, enhancing the possibility to identify disease regulating genetic variants. Here, we use a forward genetics approach in mice to identify and characterize genetic determinants of RA and related autoimmune diseases. First, we have mapped quantitative trait loci (QTL) regulating experimental arthritis using linkage analysis, and then isolated these QTL in congenic strains for in depth functional characterization. Using this approach, we make several important observations.

In the first study, we find that promoter polymorphisms regulating expression of the vitamin D receptor affect T cell activation and T cell-driven collagen-induced arthritis. These findings are particularly interesting considering the long-standing association between serum vitamin D and several autoimmune diseases.

In the second study, we discover a spontaneous insertion of a long terminal repeat which leads to a deficiency in the *SH3GL1* gene (Endophilin A2), protecting mice in several arthritis models. We are first to identify the immunomodulatory properties of SH3GL1, which may prove to be a valuable therapeutic target.

In the third study, we identify a polymorphic estrogen receptor binding site that regulates susceptibility to experimental arthritis and other autoimmune models by interfering with estrogen regulation of the T cell marker CD2. These results suggest an important role for CD2 and estrogen in shaping the sexually dimorphic immune response. Collectively, our findings make a significant contribution towards the understanding of RA genetics while demonstrating the value of animal models.

SCIENTIFIC PAPERS INCLUDED IN THIS THESIS

I. Fernandez Lahore, G, Raposo, B, Lagerquist, M, Ohlsson, C, Sabatier, P, Xu, B, ... & Holmdahl, R (2020). Vitamin D3 receptor polymorphisms regulate T cells and T cell-dependent inflammatory diseases. *Proceedings of the National Academy of Sciences*, *117*(40), 24986-24997.

II. Norin, U, Rintisch, C, Meng, L, Forster, F, Ekman, D, Tuncel, J, ... , Fernandez Lahore, G, … & Holmdahl, R (2021). Endophilin A2 deficiency protects rodents from autoimmune arthritis by modulating T cell activation. *Nature communications*, *12*(1), 1-11.

III. Fernandez Lahore, G, Förster, M, Johannesson, M, Sabatier, P, Lönnblom, E, Aoun, M, ... & Holmdahl, R (2021). Polymorphic estrogen receptor binding site causes CD2-dependent sex bias in the susceptibility to autoimmune diseases. *Nature communications.* In press.

CONTRIBUTIONS TO OTHER MANUSCRIPTS

I. Fernandez Lahore, G*, James J*, et al. Redox regulation of LAT drives T cell-mediated inflammation. Manuscript. * equal contribution.

II. James J, et al. Redox regulation of PTPN22 affects the severity of T cell dependent autoimmune inflammation. Under review in *Redox Biology*.

III. Aoun M*, Saxena A*, et al. Bone marrow-selected antigen specific regulatory B cells. Manuscript. * equal contribution.

IV. Urbonaviciute V, et al. Collagen peptide based vaccine induces T cell-mediated protection from autoimmune arthritis independent of the cuasative antigen. Manuscirpt.

CONTENTS

LIST OF ABBREVIATIONS

(m)Ab	(monoclonal) Antibody
APC	Antigen presenting cell
ACPA	Anti citrullinated protein antibody
CD	Cluster of differentiation
CFA	Complete freud's adjuvant
CIA	Collagen-induced arthritis
bCII	Bovine collagen type II
DC	Dendritic cell
E2	17-β-estradiol
EAE	Experimental autoimmune encephalomyelitis
ER	Estrogen receptor
ERBS	Estrogen receptor binding site
GWAS	Genome-wide association study
HLA	Human leucocyte antigen
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
Mbp	Mega base pairs
QTL	Quantitative trait locus
RA	Rheumatoid arthritis
RF	Rheumatoid factor
SNP	Single nucleotide polymorphism
TCR	T cell receptor
Th	T helper cell
TNF	Tumor necrosis factor
VDR	Vitamin D receptor

1 INTRODUCTION

The human is in constant contact with commensal and pathogenic microorganisms. The immune system is an intricate network of specialised cells and mechanisms that provides protection from possible infections. The immune system can be broadly subclassified in two groups, innate and adaptive.

The innate immune system is the first line of defence, comprised by a series of non-specific mechanism that include skin and mucosal barriers, complement proteins, as well as phagocytic and antigen presenting cells (APCs). Innate cells use a series of somewhat unspecific receptors to detect, and react to, shared microbial structures such as lipopolysaccharides, proteoglycans, unmethylated CpG dinucleotides, or double stranded RNA (1). Activated APCs can recruit and expand adaptive effector T cells to generate highly specific immune responses (2).

Adaptive T and B cells somatically rearrange their receptors during development. In this way, millions of T and B cell clones with unique receptor combinations are generated. To be activated, T cells require cognate antigen to be presented by APCs. B cells can be stimulated by antigen alone or in a T cell-dependent manner to produce antibodies. B cells activated in a T cell-dependent manner, however, undergo somatic hypermutation of the B cell receptor variable region, resulting in the generation of mature, highly specific antibodies. This process results in a flexible and tailored cellular and humoral immune response, but it comes at a cost. The same process can generate self-reactive clones that attack our own tissue, causing systemic or organ-specific inflammation and destruction, an autoimmune disease.

Mechanisms are in place to minimize the occurrence of autoreactive events, first at a central, then at a peripheral level. Developing bone marrow B cells and thymic T cells with autoreactive potential are deleted, receptor edited, or imparted with anergic or anti-inflammatory properties during selection (3) (4). Autoreactive cells that escape elimination face obstacles in the periphery, such as restricted tissue access and low abundance of antigen. Further, autoreactive B cells lacking essential T cell co-stimulation become anergic (5), and improperly (co)stimulated autoreactive T cells become apoptotic, anergic or develop regulatory properties (6).

However, these protective mechanisms are not perfect, and tolerance can be breached under certain circumstances. Events proposed to break tolerance include exposure of antigen from immune privileged sites after injuries (7) (8), cross reactivities between exogeneous and endogenous antigens (9), and reactivity to peripherally modified antigens (neoantigens) (10). Genetic factors are also critical. For example, rare mutations in single genes affecting central or peripheral tolerance such as *FOXP3* (11) or *AIRE* (12) can result in autoimmune conditions. Most often, however, genetic susceptibility is determined by multiple common genetic variants of low effect sizes with poorly understood and complex underlying mechanisms. Likely, it is complex interactions between predisposing genetics and environmental factors that determine the susceptibility to autoimmune diseases.

Today, autoimmune diseases can be treated with different degrees of success but cannot be cured. These diseases reduce quality of life in affected persons and are a major socioeconomic problem (13). It is imperative that we improve our knowledge to successfully establish new therapies and preventive measures. One of the most frequent autoimmune diseases, and the focus of this work, is rheumatoid arthritis.

2 LITERATURE REVIEW

2.1 RHEUMATOID ARTHRITIS

Rheumatoid arthritis (RA) is a multifactorial and chronic autoimmune disease of the joints affecting 0,5-1% of the population (14). Like many autoimmune diseases, RA is predominant in females, with a female to male ratio of approximately 3:1 (15). RA patients exhibit characteristic signs of chronic inflammation in the joints, as evidenced by synovial hyperplasia and increased signs of angiogenesis (16) (17) (18). Chronic joint inflammation promotes osteoclast formation and leads to increased bone resorption and erosion. If the inflammation is not treated, progressive destruction of the joints eventually leads to disability.

Figure 1*: Illustration of typical joint inflammation in patients with rheumatoid arthritis. Interphalangeal, metacarpophalangeal and wrist joints are most often affected. More chronic RA may also involve thickening of metacarpophalangeal joints and ulnar deviation of the fingers.*

RA can be treated with varying degree of success. Therapies focus on stopping inflammation and managing pain. Common therapeutic approaches include the use of so-called disease modifying antirheumatic drugs (methotrexate, hydroxychloroquine) in combination with non-steroidal anti-inflammatory drugs (diclofenac, ibuprofen) or glucocorticoids, as well as more recently the use fast-acting biologics such as B cell-depleting anti-CD20 or TNF blockers.

RA is diagnosed based on clinical manifestations and the characteristic presence of autoantibodies. Patients are screened for the presence of rheumatoid factors (RF) (19) and anti-citrullinated protein antibodies (ACPA) (20). RF are polyclonal and germline encoded IgM anti-IgG antibodies that often occur in the context of inflammation. ACPAs, on the other hand, are promiscuous antibodies capable of binding diverse citrullinated autoantigens (21) that are highly specific to RA (specificity up to 98%) (22).

ACPAs are not only the gold-standard for RA diagnosis but also an important stratification criterium, as not all patients have these antibodies (sensitivity up to 77%). ACPA positive and negative patients differ in e.g. disease progression (23) and genetic risk profiles (24), suggesting that RA has different underlying aetiologies. It is unclear how or whether ACPAs or RFs are disease modifiers and, in fact, ACPAs precede onset of symptoms (25). It is likely that autoimmunity develops in a yearlong subclinical phase before culminating in a chronic and destructive disease (21). Our current understanding of RA is limited in that it is derived mostly from established disease.

2.2 PATHOPHYSIOLOGY OF ESTABLISHED RA

Joint inflammation in RA is characterized by the presence leukocyte infiltrates to the joint. Local inflammation is perpetuated by the presence of effector cytokines such as IL-1, IL-6, and TNFα (26), which are secreted by, and act on, infiltrating and resident cells alike. Both adaptive and innate immune cells partake in synovial inflammation (27).

Cells of the **adaptive immune system** have generated particular interest given the characteristic presence of B cell-derived autoreactive ACPAs and the genetic risk profile in RA patients. Many of the genes associated to RA are key players in T cell and B cell activation (28), together suggesting an underlying adaptive autoimmune response. Indeed, T cell- and B cell infiltration of the joints is characteristic for RA (29), with both these cell types forming a large fraction of the synovial infiltrate (30). Infiltrating B and T cells can be found diffusely or in lymphocytic aggregates (31, 32). Here, they seem to interact productively as ACPAs evolve over time (33), suggesting a T cell-dependent B cell activation. The relevance of both of these cell types for pathogenesis is further substantiated by the clinical efficacy of T cell inhibiting CTLA4Ig (34), and B cell depleting anti-CD20 (35). Interestingly, anti-CD20 treatment does not effectively deplete antibody producing plasma cells (36), highlighting the importance of antibody-independent B cell functions. It is known that activated **B cells** are capable of secreting proinflammatory cytokines (37) and potently presenting antigen (38), thus supporting activation of synovial (autoreactive) T cells (39) and other cells. This help in activation seems reciprocal, as RA T cells also support B cell activation (40).

Synovial **T cells** in RA display a contracted T cell receptor (TCR) repertoire characterized by patient-individual clonal expansion (41), some of these clones likely recognizing citrullinated autoantigens (42). Expanded synovial T cells take on a terminally differentiated and activated phenotype (40) (43), losing expression of costimulatory molecules such as CD28, while at the same type gaining other receptors that permit activation in an antigen-independent manner (44). T cells from RA synovia show increased signs of replicative stress (45) and are, to some extent, resistant to apoptosis (46). These cells potently produce IFN-γ, and to some extent IL-17 (47–50), enhancing inflammation (51) and joint destruction, for example by promoting osteoclastogenesis (52). In fact, the frequency of Th1 type cells (producers of IFN-γ) correlate with disease activity (53). Further insight can be extracted from animal models of RA, which demonstrate that transfer of arthritogenic T cells is enough to elicit the development of autoimmune arthritis in naïve recipients (54). Thus, although details are still to be sorted out, current evidence strongly suggests that T cells are important drivers of the autoimmune response in RA. Therefore, the study of T cell modifiers remains central to understanding RA pathogenesis.

Perhaps as a consequence to an initially T cell driven autoimmune response, innate type cells are recruited to the joint. Resident and infiltrating **innate immune cells** are activated by local accumulation of proinflammatory cytokines and immune complexes, producing cytokines, chemokines, prostaglandins, and leukotrienes in response. Activated **neutrophils** furthermore have the capacity to produce high levels of reactive oxygen species and secrete nuclear extracellular traps to promote inflammation (55). **Macrophages** are present in increased numbers in the RA synovial membrane and are believed to contribute to inflammation by presenting antigen and producing cytokines such as TNFα, IL6, IL-1β, IL-12 and IL-23 (56).

This recruits monocytes and neutrophils, and has repercussions for osteoclast formation, and T cell and fibroblast activation. However, recent studies also suggest that resident synovial membrane macrophages originally serve to shield and protect articular structures from infiltrating cells (57). **Mast cells** are also expanded in the arthritic synovial membrane, where they secrete proteinases (58) and inflammatory mediators (59). Interestingly, they can promote T cell responses by presenting antigen (60) but are also capable IL-17 producers themselves (61).

Joint inflammation is not only driven by immune cells, but also by resident cells of **mesenchymal** origin. Activated synovial lining **fibroblasts** are imprinted with an aggressive phenotype that perseveres for several months of culture, producing proinflammatory cytokines, chemokines, prostaglandins and matrix-degrading enzymes (62). Similarly, **chondrocytes** are activated by IL-1 and TNFα to secrete proteolytic enzymes, leading to extracellular matrix degradation (63).

Evidently, the rheumatic joint consists of a highly complicated network of proinflammatory mediators. A recurring theme is the response to -and secretion of- proinflammatory cytokines, in particular IL-1, IL-6, and TNFα. All these cytokines can be readily detected in RA synovia. Accordingly, targeting of these pathways presents one of the best therapeutic options in RA alongside anti-CD20. While current therapeutic approaches successfully manage synovitis, an estimated 5-20% of patients do not respond well to available treatments (64). Additionally, biologics carry important added risks such increased susceptibility to infections or malignancies (65). To develop new targeted therapies, it is paramount to understand which factors predispose an individual to develop RA.

2.3 PREDISPOSING FACTORS IN RA

Studies have estimated RA heritability at around 50 % (66). The number is not precise, as estimates vary between 12-68%, but it is clear that RA has a significant genetic component. On the other hand, this also means that there must be a significant environmental contribution to RA susceptibility. We will first briefly cover commonly suggested environmental factors to then dedicate more time to RA genetics.

2.3.1 Predisposing environmental factors

Several environmental factors have been associated to RA. Some of the most frequently mentioned examples include cigarette smoking (67), infections (68) (69), vitamin D status (70), use of oral contraceptives (71), and exposure to mineral oils (72).

It has been suggested that infections may trigger cross-reactivities to self through molecular mimicry. For example, Epstein-Barr (69) viral proteins exhibit similarity to HLA-DR4 (73) (74) (75). Similarly, smoking (67) and infection with *Porphyromonas gingivalis* (68), are suspected to elicit autoimmune responses by promoting the formation of neo-antigens. Special attention has been given to neo-antigen formation as a result of hypercitrullination, given a predisposition of RA-associated HLA alleles to bind citrullinated proteins (76), and the association between polymorphisms of the citrullinating enzyme *PADI4* (77) and RA. Both smoking (78) (79) and periodontitis (80) are suspected to promote citrullination, and fittingly associate with ACPA seropositive RA (81).

A particularly interesting association in the context of study I is that of RA and vitamin D status. Several studies have reported an association between low serum vitamin D and incidence of RA (82) (83) (84). In study I we demonstrate that the genetic predisposition to over and under express the vitamin D receptor affects the immune response in experimental models of autoimmunity.

Similarly, in study III we provide evidence that may help understand associations such as the one between RA and use of oral contraceptives (85) (86). A key role for sex hormones in RA susceptibility has long been suspected, and this notion is further supported by the female preponderance in RA (87), as well as studies reporting disease modifying effects of hormone replacement therapy (88) (89). In study III we describe how sex hormones such as estrogen may directly contribute to the development of RA by regulating the expression of the T cell costimulatory molecule CD2.

While our studies provide some insight into possible mechanisms fuelling some of the environmental associations identified to date, in general, associative studies investigating environmental factors in RA have had limited success. Environmental factors may act years before RA onset and are thus difficult to identify. Perhaps because of this, studies investigating environmental factors are scarce and often lack reproducibility (90). In contrast, more energy has been spent understanding the genetic determinants of RA.

2.3.2 Predisposing genetic factors

The estimated degree of heritability in RA varies depending on the method used. Comparison of disease discordance between monozygotic and dizygotic twins found heritability estimates of 12-60%, familial aggregation studies of 50%, and methods considering the cumulative influence of all SNPs on RA phenotypic variance suggest a heritability of 19-52% (66). While the heritability estimate may not be exact, it is clear that the contribution of genetics to RA is significant.

Already in the late 1970s it was observed that PBMCs from RA patients are unusually compatible with each other in mixed lymphocyte cultures (91), and it was quickly recognized that this was due to similarities in the *HLA* alleles of RA patients (92). This association was soon refined to the *HLA-DRB1* gene and remains the strongest genetic association to RA to date (93), accounting for around 30% of the total estimated heritability (94). It has been proposed that several of the *HLA-DRB1* susceptibility alleles share a common amino acid sequence in the third hypervariable region of DRβ1 (shared epitope hypothesis (92)). This shared motif may favour binding of arthritogenic peptides although a main candidate remains to be singled out.

The emergence of high throughput sequencing and genotyping methods has enabled genome wide association studies that helped clarify the genetic landscape of RA by further identifying hundreds of non-MHC risk loci. Interestingly, many of the non-MHC associations are to genes relevant to T cell biology and activation, further cementing the importance of T cells for RA pathology. Some of the most cited findings include *PTPN22, STAT4, TRAF1-C5, IL2RA, CD28,* and *CTLA4* (95), as well as more recently Th17-related *CD26* and *CCR6* (96) among others. T cell-unrelated genes with other relevant functions, such as the citrullinating enzyme *PADI4* (95), have also been detected. Some of these associations can be ethnic group-specific,

such as is the case for *HLA-DRB1* alleles or for Caucasian-specific *PTPN22*, but most reproducible associations are shared across populations (97). In contrast, seropositive and seronegative RA seem to have a different genetic risk factors (24). For example, *HLA-DRB1* and *PTPN22* seem to associate primarily with seropositive RA (98) (99).

Despite the considerable progress made by modern genetic association studies, it should be noted that the newfound loci only make small individual contributions to explaining the overall heritability of RA (figure 2). Cumulatively, the current genetic findings explain an estimated 20% of the total phenotypic variance (97) as opposed to approximately 40% estimated by familial and twin studies. This gap has been appropriately termed *missing heritability* and remains an important gap in the genetic understanding of RA.

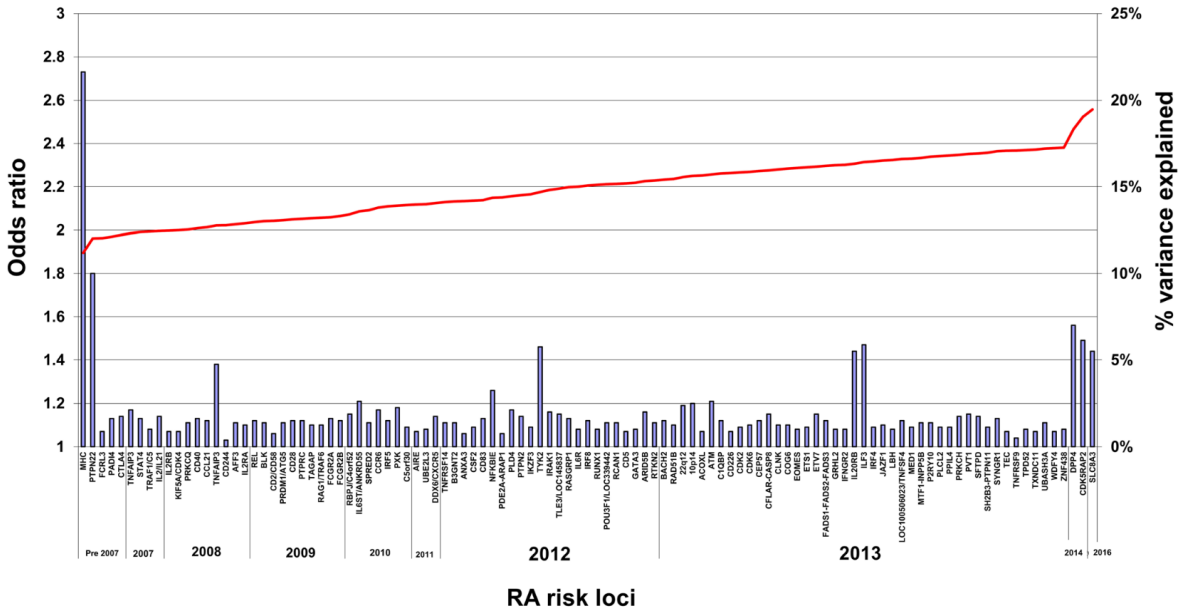

Figure 2. *Cumulative proportion of the total phenotypic variance explained by findings of current genome wide association studies. Adapted from Viatte and Barton, 2017* (66). *CC-BY 4.0.*

2.3.3 Limitations of human genetic studies

The missing heritability can be explained by an over-estimation of the heritability in twin studies due to limited sample numbers, but also by failure of modern genome wide association studies to capture all genetic variation due to technical limitations. For example, genome wide scans are designed to detect common variants and are blind to low-frequency variants (100). This is problematic because RA likely cannot be explained only through common genetic variation (101). Further, key but structurally complex genetic regions can be excluded from analysis due to technical difficulties (102). It should also be considered that genotyping markers may be in incomplete linkage disequilibrium with causative variants (103). Also problematic is the analysis of higher order interactions to be able to account for the likely very relevant contribution of gene-gene (epistasis) and gene-environment interactions. Work from model organisms such a *Drosophila melanogaster* has made clear that complex traits are shaped by many genetic variants of moderate effect size that act in concert with each other and with the environment (104).

Last but not least, human genetic association studies also suffer from statistical limitations. They lack sensitivity towards loci of weak effect size (105) given the heterogeneity in the human population and high statistical thresholds imposed by multiple testing issues (106) (107). As a result, it is likely that several risk loci remain to be identified. Figure 3 visualizes the astounding differences in power when trying to detect strong versus moderate-to-weak RA susceptibility loci. This is unfortunate, as we demonstrate in the present work that polymorphisms in genes with moderate effect size are likely critical co-determinants of disease development and perpetuation. We find that polymorphisms in genes such as *VDR*, *CD2*, or *SH3GL1* can be critical to fine-tuning T cell responses and inflammation.

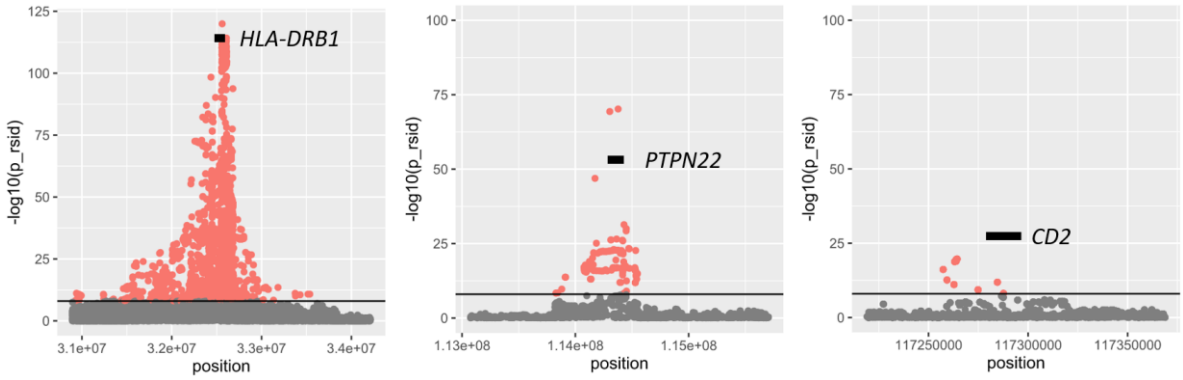

Figure 3: *Regional association plots showing strength of association with RA across the HLA-DRB1, PTPN22, and CD2 locus. Black bars give an approximation of the respective gene. The genome-wide significance threshold is indicated by the lower black line. GWAS summary statistics were extracted from the IEU open GWAS project* (108).

Once a candidate risk locus is identified, further resistance comes in the form of functional interpretation. Unfortunately, most identified variants in genetic association studies concentrate in non-coding regions (109) (probably of regulatory nature (110)) and are thus difficult to interpret. Often, non-coding variants have regulatory properties, affecting regulation of nearby genes (111) (112). However, clarifying these mechanisms is time and resource consuming, a process further impeded by cell type or general context specificity of the effects of some variants (113). Additional difficulty is given by the fact that a sequenced SNP can be in high linkage disequilibrium to actual causative variants.

Evidently, disentangling the genetic component of RA is complicated. Complementary approaches to human genetic association studies are needed to broaden our understanding of genetic variants contributing to disease susceptibility. Using animal models of arthritis, we demonstrate in study I and study III how non-coding variants can control gene expression to regulate immune phenotypes.

2.4 USING MOUSE MODELS TO STUDY RA GENETICS

Contrary to human samples, work with animal models provides stable experimental conditions, genetic uniformity, accessibility to disease-relevant tissues, and a broad genetic toolkit for functional characterization of candidate genes.

Mice and human shared a common ancestor 60-80 million years ago, but identity and arrangement of genes have largely stayed the same (114). The mouse and human genome are

highly conserved (115). This is also true for immune biology, although it should be recognized that discrete differences exist between species (116) (117) (118). Thus, rodents are an excellent model system to identify and characterize genetic factors regulating immune, as well as other, traits. In theory, it is possible to dissect the genetic component of a complex trait using only a few hundred mice (119), whereas in humans the same effort may require tens of thousands of individuals (120). Optimally, of course, human and mouse studies can be integrated, with findings from each species informing the biology of the other. To generate relevant and translatable findings, mouse genetic studies require well characterised model systems closely resembling the human trait or pathological condition under study.

2.4.1 Inducing RA-like arthritis in mice

Several models exist to induce RA-like autoimmune arthritis in mice and rats. Many of these models consist in immunizing mice with common cartilage proteins in adjuvant. Some examples include collagen type II (121) and type XI (122), as well as cartilage oligomeric matrix protein (123). Nevertheless, it is also possible to induce arthritis using systemically expressed antigens such as glucose-6-phosphate isomerase (124) or using mineral oils, such as is the case in pristane-induced arthritis (125). This last model is particularly interesting considering the association between RA and exposure to mineral oils (72). Not all arthritis models need to be induced. Mice with transgenic overexpression of TNF-α (126) or mice with a spontaneous mutation in the T cell signalling ZAP-70 protein (127) develop spontaneous arthritis.

In the present work, we have mostly employed the collagen-induced arthritis (CIA) model (121), as it is a well characterized and widely used rodent model of RA with several parallels to the human autoimmune condition. In this model, an autoimmune arthritis is triggered by immunizing mice with heterologous bovine collagen type II (bCII) in Complete Freund's Adjuvant (128). This elicits an autoimmune response driven by CII autoreactive T cells (129) (130) (131) and characterized by the development of arthritogenic autoantibodies (132). A typical CIA experiment is depicted in figure 4. Much like in RA, CIA mice exhibit progressive and severe joint lesions with T cell infiltrates (133), cartilage injury, bone remodelling (134), and synovial hyperplasia (135). Similarly, several key RA cytokines are also involved in CIA immune pathology, including IL-17 (136), IFN-γ (137) , IL-6 (138), TNFα (139), or IL-1β (140).

CIA is a convenient screening model, as it is well characterized, incidence is high in susceptible strains, and arthritis development can be easily monitored by scoring visible inflammation of front and hind paws. Using this model, our research group has set out to identify genetic determinants of autoimmune arthritis in a forward genetics-based approach. 'Forward' in this case describes the direction of the workflow, starting from a phenotype and working towards the unknown underlying genetic mechanism. Importantly, this method is hypothesis-free.

2.4.2 Mapping quantitative trait loci in the mouse

The positional identification of risk genes consists of two phases. First, genetic loci affecting a quantitative phenotype of interest (quantitative trait loci, QTL) are roughly mapped by linkage analysis, or in genetic association studies. Then, the identified QTL are isolated in congenic

strains to enable functional studies with the aim of identifying the underlying genetic mechanisms. This process is illustrated in figure 5.

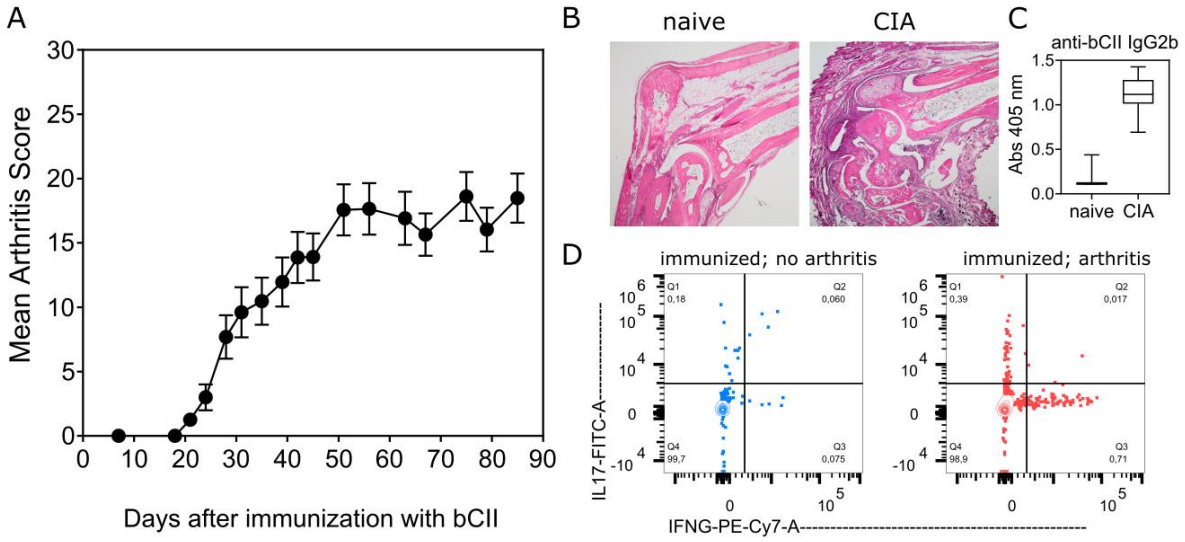

***Figure 4: Collagen-induced arthritis**. A) Typical CIA diseases course in C57BL/10.H2^R mice. Data summarized as mean (SEM) from n = 15 mice. B) Representative H&E stain of tibia and talus in healthy (left) and CIA mice (right). Note increased inflammatory cell infiltrates (violet cell nuclei), thickening of synovial layers, and bone erosion in CIA mice. Final magnification 25 X. C) ELISA showing development of anti-bCII IgG2b antibodies in CIA mice from (A) (day 35). D) Flow cytometry plots showing IL-17 and IFN-γ production in CD4$^+$ lymph node T cells from CIA mice after in vitro recall with bCII. The dot plot on the left (blue) is from an immunized mouse that did not develop arthritis, and the one on the right (red) from a mouse that developed visible arthritis. bCII, bovine collagen type II.*

Congenic strains are created by intercrossing a donor with a recipient strain, and then backcrossing the offspring to the recipient strain over several generations while selecting for the genetic locus of interest. Once done, a congenic mouse has the donor allele on a single, selected genomic locus on an otherwise recipient background. In this way, an arthritis susceptibility locus can be isolated on an arthritis resistant background or vice versa. Established congenic mouse strains allow detailed study of a QTLs in a genetically controlled environment. Such strains have for example been used in pioneering work to demonstrate the importance of the major histocompatibility H-2 locus in the development of adaptive immune responses (141).

Much like in human genetic studies, the study of QTL in animals is not straightforward. QTL identified from F2 crosses are usually large, often containing tens to hundreds of polymorphic genes that differ between parental recipient and donor strains. In its simplest form, QTL resolution can be increased by further backcrossing to parental strains or further (advanced) intercrossing of F2 mice (142, 143). The interval of interest can be isolated in interval-specific congenic strains that are further intercrossed. Recombinants within the interval of interest are then selected and phenotyped. This procedure is time and resource consuming but can result in small 1-2 Mbp long fragments that contain 10-20 polymorphic genes. To make achieving single gene resolution more feasible, alternative approaches have been developed, each with their own challenges.

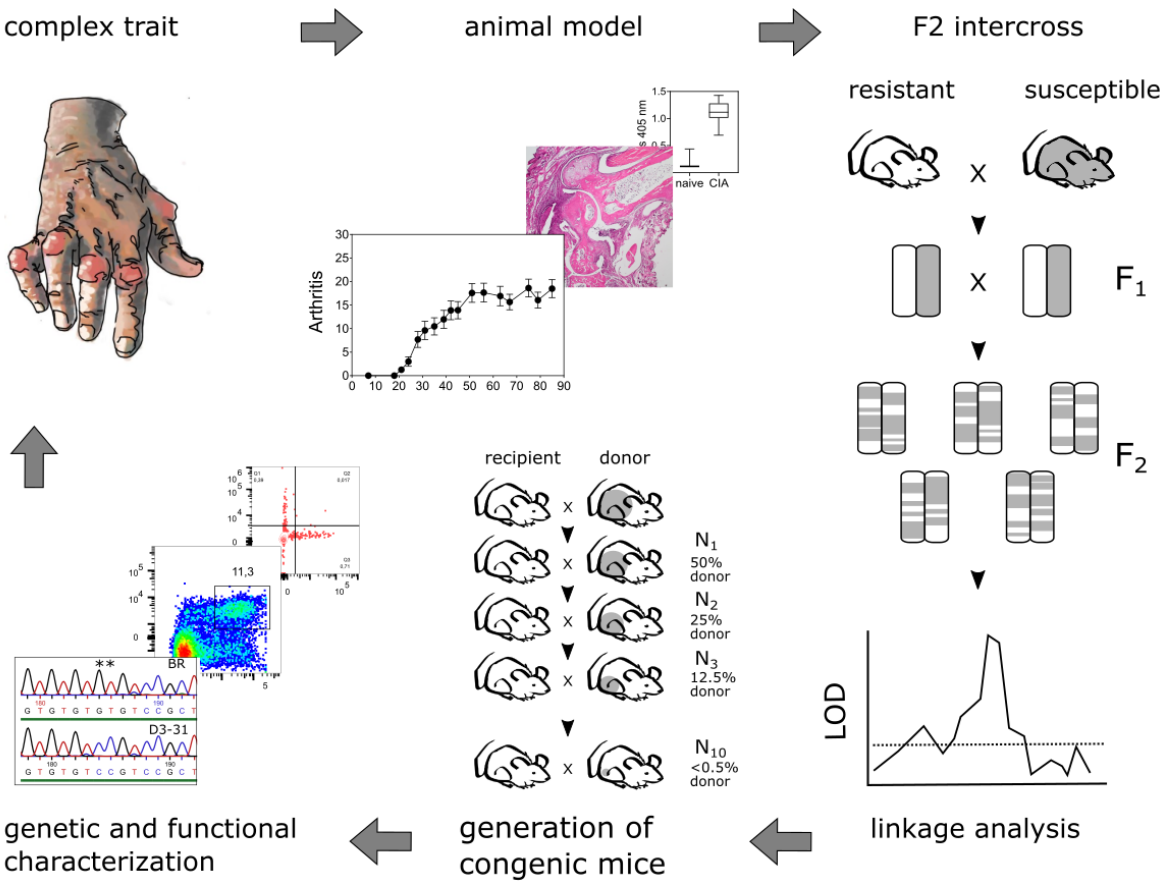

Figure 5. Positional identification of genes regulating complex traits using mouse models. *A suitable mouse model is selected to approximate human disease, in this case CIA. For linkage analysis, arthritis susceptible (donor) and resistant (recipient) strains are crossed, and the resulting F2 mice are individually genotyped and phenotypically evaluated. Using computational analysis, one can establish which genetic markers segregate with the phenotype of interest. The logarithm of odds (LOD) score estimates the concordance between a genetic marker and a phenotype. F2 mice of interest are then backcrossed to a parental strain to generate congenic lines. Repeated backcrossing progressively reduces the overall genetic contamination, halving the genetic content of the donor each generation. After 10 rounds of backcrossing, the congenic strain differs only at a defined genetic locus from the recipient parental strain. This locus can be further refined by backcrossing or intercrossing* (144). *The resulting congenic mice are used for further genetic and functional analysis.*

One such example are recombinant inbred lines (RIL). Here, F2 mice are randomly intercrossed over several generations to accumulate recombinations, and the resulting mice then sibling mated for enough generations to generate identical offspring. Importantly, the resulting commercially available lines need only be characterised once on a genotypic level and can then be screened for several phenotypes of interest. The collaborative cross project (145) builds on the idea of RIL but uses 8 (instead of the conventional 2) founder strains carefully selected to maximize genetic diversity. The use of CC mice could potentially result in QTL resolution of approximately 8 Mbp (146). This resolution can however be further improved by the use of complementary methods such as yin-yang crosses (147). The limiting factor in the use of RI mice is the amount of mice required to resolve phenotypes and the attached cost and logistics

of maintaining thousands of lines. Furthermore, RIL often are bad breeders due to extensive inbreeding.

Breeding problems can be circumvented by using outbred heterogeneous stock (HS) mice (148)(149). HS mice result from advanced intercross of several inbred lines bred in pseudo-random manner to minimize loss of genetic variability. Recombinations accumulate over generations so that it becomes technically feasible to map QTL with sub-Mbp accuracy. Every mouse is unique and must be individually genotyped in a genome wide scan much like in human studies. The disadvantages of genome-wide analysis using HS mice lies in the complexities of the analysis, need for large cohorts (thousands of mice), high-density genotyping required (>6000 markers), possibility of false positives, and multiple testing issues.

Finally, it is possible to avoid breeding work altogether by mapping QTL *in silico*. Since common inbred mouse strains have shared ancestry, it is possible to identify regions identical by descent affecting quantitative traits. Here, statistical power can be low, as there is only a limited number of inbred lines available, and assumptions need to be made regarding conservation and inheritance pattern of identified haplotype blocks between inbred strains and original founders. Further, a general limitation from genetic associations studies are false positives (150). Regardless of the method used for mapping QTL, ultimately linkage analysis and further functional studies are needed to conclusively prove findings from association studies.

2.4.3 Arriving from QTL to causative gene

Distilling a causative genetic variant from a QTL is a challenging endeavour (151) (152). Usually, the driving genes or variants in a QTL cannot be exactly positioned but instead need to be inferred. For example, polymorphic genes within a QTL may be prioritized if they are immediately relevant to the investigated trait, especially in the presence of non-synonymous polymorphisms (153) (154) (155). Likewise of interest, may be polymorphic genes exhibiting altered expression levels (156), for example as a consequence of non-coding polymorphisms in promoters or enhancers. Once a candidate gene is singled out, the hypothesis is tested by complementary approaches, to demonstrate functional relevance of the proposed candidate.

Popular strategies to demonstrate functional relevance include knockout, inhibition, or overexpression of the candidate transcript in a relevant setting (157), as well as stimulation or inhibition of the target protein using agonists or antagonists, respectively. A particularly elegant method is quantitative complementation, where congenic mice are crossed to a null allele of the candidate gene to compare the ability of the congenic and wildtype alleles to compensate for the deficiency of the gene under investigation (158). Finally, the identified pathway should also be explored in humans to corroborate overall relevance of the findings obtained in mice.

While systematic mutagenesis or systematic knockout may seem like more straightforward alternatives to genetic mapping studies, these approaches have their own limitations (159). An approximated 15% of genes are essential and are thus not optimal target for knockout strategies. Further, knockouts approaches result in complete loss of gene function, which does not reflect the physiological differences generated by natural variation. As a result, knockouts may elicit compensatory mechanisms or disrupt important genetic or protein interactions, resulting in artificial phenotypes that may not mirror the effects of common genetic variation. Therefore,

established congenic strains and linkage analysis remain a valuable tool as they not only allow the conclusive identification of important genetic determinants, but also provide a physiological context to study biological mechanisms in a meaningful manner.

3 RESEARCH AIMS

The genetic component of complex autoimmune diseases such as RA remains poorly understood. To aid in this process, we previously set out to identify QTLs critical to the development of autoimmune arthritis in mouse and rat model systems. The aim of this study was to identify and characterize the genetic polymorphisms and molecular pathways underlying several previously reported QTL.

Specifically, the aims were as follows:

- Identify the genetic and molecular mechanisms driving Cia37 (142)
- Identify the genetic and molecular mechanisms driving Cia21 (160)
- Establish the molecular mechanisms underlying the spontaneous Pia34 QTL (161)

4 METHODS

4.1 A brief overview of the animal models applied in this study

Collagen-induced arthritis. CIA is induced by immunizing mice intradermally with heterologous collagen type II in CFA. Susceptible mice develop joint inflammation characterised by synovitis, pannus formation, and cartilage and bone damage. CIA is MHC, T cell, and B cell dependent. This model is described in more detail under 2.4.1.

Pristane-induced arthritis. Rats immunized with a single intradermal pristane injection develop severe and chronic arthritis with symmetrical joint inflammation. PIA is MHC and T cell dependent and indeed T cells infiltrates can be found in PIA joints. However, the exact mechanism remains unclear as possible immunodominant T cell or B cell epitopes are not known. In contrast to CIA, B cells are considered less important in PIA because serum does not transfer disease (162).

Glucose-6-phosphate isomerase-induced arthritis. Transgenic K/BxN mice develop antibodies against the GPI protein, causing spontaneous arthritis. GIA can also be induced by immunizing mice with GPI full protein or immunodominant peptide hGPI325-339. Mice develop a monophasic arthritis that is T cell and B cell dependent (163). Anti-GPI autoantibodies can be found in a subset of RA patients (164).

Collagen antibody-induced arthritis. Arthritis is induced through intraperitoneal injection of a cocktail of monoclonal arthritogenic antibodies targeting different epitope of the collagen type II protein. Mice develop a short lived monophasic arthritis that is mediated by neutrophils and macrophages activated by the presence of antibody aggregates on the cartilage. In this context, complement and FcγRs seem to play important roles. CAIA is T and B cell independent, and a useful model to study the effector phase of CIA (165).

Delayed-type hypersensitivity. The DTH model is, in essence, an in vivo antigen recall assay driven by Th1-type cells infiltrating the skin (166). Mice are sensitized by immunizing with antigen in CFA, and antigen-specific T cells are allowed to expand for 10 days. Thereafter, mice are challenged with antigen intradermally and tissue swelling is measured after 24-72hrs using a calliper to evaluate T cell responses to the antigen.

Experimental autoimmune encephalomyelitis. EAE is a model of multiple sclerosis. Mice are immunized intradermally with myelin derived antigen emulsified in CFA. Pertussis toxin can be administered to increase susceptibility. Mice can be immunized with myelin oligodendrocyte protein, proteolipid protein, myelin basic protein, or spinal cord homogenate. The result is a progressive, ascending paralysis of the extremities that begins at the tail. Paralysis is caused by monocytes and T cells infiltrating the central nervous system and causing axonal demyelination. $CD4^+$ T cells are central for EAE with contributions from Th1 and Th17-type cells chiefly producing IFN-γ, IL-17, and GM-CSF. Both myelin-reactive $CD4^+$ and $CD8^+$ T cells can transfer EAE, and are required to induce it. B cell-derived autoantibodies seem less critical, with B cell acting mainly to support T cell responses (167).

DSS-induced colitis. DSS-induced colitis is a quick, 10 day model of inflammatory bowel disease. Administration of dextran sodium sulfate (DSS) in the drinking water induces severe colitis in susceptible mice by damaging the intestinal epithelium. Disease activity is evaluated by following rectal bleeding, weight loss, and stool consistency. This model is driven by innate cells like neutrophils and monocytes, and is T and B cell independent (168).

5 RESULTS AND DISCUSSION

5.1 Study I - Vitamin D receptor polymorphisms regulate T cells and T cell-dependent inflammatory diseases.

In this study we positionally identify polymorphisms in the vitamin D receptor (VDR) promoter as regulators of T cells in autoimmunity. We find that mice that overexpress VDR in activated autoreactive T cells develop more severe autoimmune arthritis and that physiological levels of vitamin D are required for normal T cell biology and activation. Due to the physiological nature of our model, we can rule out calcaemic confounding factors. We also show that lymphocytes in rheumatic joints highly upregulate VDR, suggesting that our finding are relevant to human RA. In conclusion, our results suggest an inflammatory role for VDR in activated T cells.

The immunomodulatory properties of vitamin D were first recognized when it became evident that cod liver oil, which is high in vitamin D, arrested development of tuberculosis (169) (170). Further studies then showed that vitamin D availability is critical to innate immune responses and bacterial clearance. At the same time, however, vitamin D deficiency has been associated to several autoimmune conditions (171) (172), and together with in vitro studies fuelled the idea that vitamin D may have a key immune dampening -rather than activating- role during adaptive responses (173) (174) (175) (176). Naturally, the concept of vitamin D as a dietary and inexpensive supplement to help restrain inflammation generated great interest.

However, whether vitamin D has meaningful immunoregulatory properties in a biological context has remained questionable, as the beneficiary effects of vitamin D supplementation in double-blinded clinical trials are not entirely convincing (177) (178) (179). Moreover, studies often demonstrate immunomodulatory properties of vitamin D using supraphysiological doses that provoke undesired calcaemic side effects. Our study brings a unique perspective to this issue by exploiting naturally conserved genetic variation to investigate the immunomodulatory effects of vitamin D and VDR in a physiological setting. Our findings indicate that VDR in a physiological range supports T cell activation, and, in a broader sense, help cement the importance of vitamin D for immune system biology.

5.2 Study II - Endophilin A2 deficiency protects rodents from autoimmune arthritis by modulating T cell activation.

This work originated from the observation that a subset of rats in the otherwise highly arthritis-susceptible DA strain inexplicably became resistant to arthritis. Using linkage analysis, we found this protective phenotype to be driven by a QTL on chromosome 9 (Pia43). Sequence analysis revealed a spontaneous insertion of a viral long terminal repeat in the first intron of the *SH3GL1* gene, resulting in silenced transcription of *SH3GL1*. We speculated that *SH3GL1* deficiency was protecting rats from arthritis, and indeed could confirm this observation in *SH3GL1* knockout mice. Functional characterization revealed that SH3GL1 affected T cell activation by interfering with TCR internalization. Finally, we showed that *SH3GL1* is overexpressed in PBMCs from RA patients, suggesting a role for SH3GL1 in the pathogenesis of autoimmune arthritis.

Genetic and immunological studies strongly suggest that aberrant activation of autoreactive T cells is key to the chronicity and lack of resolution in autoimmune diseases. Although

autoreactive T cells are promising therapeutic targets, development of (antigen-specific) T cell-based immunotherapies has proven challenging. Monoclonal antibody-based targeting of broad T cell populations (180) has had limited success, and more specific strategies are either still under development (vaccination-type strategies) (181) or have proven ineffective (oral type II collagen) (182) (183). Thus, there is a need to broaden our understanding of T cell regulation with the aim of improving T cell-based therapies. Our results contribute to this by identifying SH3GL1 as a key regulator of T cells and a promising new therapeutic pathway to target (autoreactive) T cells.

5.3 Study III - Polymorphic estrogen receptor binding site causes CD2-dependent sex bias in the susceptibility to autoimmune diseases

In this study we identify a polymorphic estrogen receptor binding site (ERBS) leading to sex-dependent differences in several models of autoimmunity. This ERBS orchestrated expression of surrounding genes in an estrogen-dependent manner, including the T cell costimulatory molecule CD2. As a result, polymorphisms in this ERBS interfered with hormone-dependent regulation of gene expression. Sex-specific changes in CD2 expression then culminated in sexually dimorphic T cell responses. Our findings have several implications for human autoimmunity.

First, we provide functional evidence that polymorphisms in the CD2 locus are relevant to RA susceptibility. The CD2 locus has been associated to RA before (184, 185), however association data is limited and functional evidence lacking. We demonstrate that polymorphisms in the CD2 locus are important modulators of autoreactive T cell responses and autoimmune disease in mice. Further, we show that CD2 expression is increase in RA synovia and that CD2 expression correlates with disease activity, suggesting that CD2 polymorphisms affecting CD2 expression are likely relevant to the development or perpetuation of human autoimmunity.

Secondly, this study gives insight into the mechanisms driving sex differences in the immune response. RA not only occurs predominantly in females, but its symptoms often recede during pregnancy (186) (187) and have even been coupled to the menstrual cycle (188). Further, oral contraceptives (85) (86) and HRT (88) associate positively with RA (although contradictory reports exist (189)). Together, these data suggest an important immune modulatory function of sex hormones. Indeed, both $CD4^+$ T cell and humoral responses are sex biased (190) and often stronger in women. ERs are expressed in $CD4^+$ T cells (191), and likely play important regulatory functions during immune responses. This is supported by studies showing that E2 promotes T cell responses at physiological concentrations (192) (193) (194). In the present study, we show that E2 shapes the sexually dimorphic immune response by regulating the expression of the T cell activation marker CD2.

Finally, this study also demonstrates the importance of sex-genotype interactions for regulating (auto)immune responses. Sex-genotype interactions are an attractive mechanism of action to explain the sexual dimorphism in autoimmunity, but it has been challenging to collect conclusive evidence beyond associative findings. Our study not only shows that genotype-sex interactions shape (auto)immune phenotypes, but also provides a tangible mechanism by demonstrating how polymorphisms can act in a sex-specific manner by interfering with binding of key hormone receptors.

6 CONCLUDING REMARKS

There is a clear need to further understand the substantial genetic component in RA and in other autoimmune diseases. Technological advances in sequencing and high throughput genotyping made the full characterization of complex genetic diseases such as RA a tangible goal. However, this has been bottlenecked by unexpected degrees of complexity. While the expectation to genetic associations studies was to find a discrete number of moderate effect size-QTL with single underlying genes, complex diseases instead turned out to be governed by a sea of difficult to interpret and interacting low effect size QTL (<5%) (146) (160) (195), with only a few dominant QTL of strong effect size. A reason for this structure may be high selective pressure on variants with large effect sizes.

Cloning QTL with small effect size is fundamentally problematic because it means identifying and understanding small and interweaving molecular and genetic effects. The problem of understanding complex traits is, at its core, the problem of understanding small genetic effects. This is only aggravated by genotype-environment interactions. To successfully study complex traits, model systems are required that reduce environmental and genetic variability while enabling thorough functional analysis. Mouse models meet these requirements while accurately recapitulating human (patho)physiology. Even in mouse models, however, achieving single gene resolution from large QTLs is laborious. QTNs and causative genes need not only be identified, but also functionally interpreted. Only a few studies using congenic strains have been able to achieve this (155) (196). Historically, cloned QTL have had exceptional effect sizes. Of over 2000 reported QTL, only 1% have arrived at a single gene (146).

The results we present here are the culmination of more than a decade of successful genetic work to identify and characterize key genetic determinants of autoimmune arthritis in mice. Our findings aid in our understanding of the genetic architecture of complex autoimmune diseases, providing functional evidence for the relevance of non-coding genetic variants in the development of complex traits. The importance of our findings is not limited to RA, as complex autoimmune diseases share many mechanistic features. This study thus helps clarify recurring questions in the field of inflammation and autoimmunity, such as the contribution of vitamin D receptor polymorphisms (197) or the importance of sex-genotype interactions for the development and perpetuation of autoimmune diseases (198). Additionally, we identify a novel immune regulator which may prove to be therapeutically useful (199).

7 ACKNOWLEDGEMENTS

I would like to start by thanking my main supervisor **Rikard Holmdahl** for the opportunity to do my PhD in his research group and for being optimistic in every situation. I will dearly miss your hands-off supervision style.

I also have to thank my co-supervisor **Kutty Selva Nandakumar** for originally recruiting me as a PhD student and securing initial funding. As well as my other co-supervisor **Liselotte Bäckdahl** for being generally supportive.

Further, I would like to thank colleagues that shaped me on a scientific level: **Bruno R** for teaching me the basics of experimental immunology, providing the groundwork for the VDR project, and guidance during manuscript writing. **Michael F** for doing important groundwork in the CD2/ER project and providing advice when writing the manuscript. **Martina J** for introducing me to genetics. **Jonathan T** and **Sabrina H** for being excellent scientist role models. **Min Y** for teaching me to be disciplined. **Johan B** for teaching me to be thorough and use common sense. **Simon G** for teaching me to focus. **Vilma U** for showing me that good science also needs time and patience. **Amit S** for teaching me to spend more time thinking. **Susanne W** for teaching me to be critical of everything. **Changrong G** for inspiring me to dive into bioinformatics. **Bingze X** for inspiring me to be positive and work harder.

On a more personal note, I would like to thank: **Mike A** and **Clara MH** for being good friends and excellent companions during my PhD journey. **Jaime J** for being in many ways like me. **Yibo H** for fun breaks and sharing my passion for technology. **Ulrika N** for all the fun and your ability to make the day-to-day a rollercoaster. **Alexander K** for overshadowing everything with your offbeat personality. **Michael B** for not mincing words and uncompromisingly defending your interests. **Ana C** for sharing my views on fashion and taste in mediocre pop music. **Laura RC** for gossips in Spanish and your endless and contagious energy to pursue extracurricular activities no matter how hard the work day. **Erik L** for having memorized -and being able to recite- seemingly all existing epidemiological literature. **Danielle V** for your unapologetic laugh. **Angel YM** for everything, from warmly welcoming me when I arrived, to always fixing my salary during my prolonged studies. **Zhongwei X** for being a friendly person and troubleshooting everything. **Lei C** for being the politest person. **Weiwei C** for funny and unexpected comebacks. **Huqiao L** for always being professional and willing to help. **Xiaojie C** for saying things how they are. **Taotao L** for keeping the lab in shape and being my biggest fan. **Biborka BV** for fun moments and your quirky character. **Outi S** for being cheerful and laughing at my (mediocre) comments. **Betül T** for your love for mechanical engineering. **Zeynep S** for your deadpan humour. **Rajan KP** for being a nice guy and motivated postdoc. **Jianghong Z** for showing it to everyone. **Meng L** for being friendly and always smiling. **Guanzhi** for your cool style. **Carlos P** for bringing closer my South American home. **Pierre S** for being forthcoming, helping with proteomics, and being a great tennis partner. **Christian B** for constant wittiness. **Amirata SD** for having anecdotes to every topic. **Magali, Parisa, Arian and Oksana** for providing entertainment and laughs to balance work over many dinners. **Aitor, Maria, Caitrin, Sunjay, Fadwa, Amo, and Yildiz** for being a first surrogate family when I arrived in Sweden on a dark October.

And finally, I thank **my family**, for continuous support in everything and bringing much needed light-heartedness to every situation.

8 REFERENCES

1. S. E. Turvey, D. H. Broide, Innate immunity. *Journal of Allergy and Clinical Immunology* (2010) https:/doi.org/10.1016/j.jaci.2009.07.016.
2. S. J. Gaudino, P. Kumar, Cross-talk between antigen presenting cells and T cells impacts intestinal homeostasis, bacterial infections, and tumorigenesis. *Frontiers in Immunology* (2019) https:/doi.org/10.3389/fimmu.2019.00360.
3. L. Klein, E. A. Robey, C. S. Hsieh, Central CD4 + T cell tolerance: deletion versus regulatory T cell differentiation. *Nature Reviews Immunology* (2019) https:/doi.org/10.1038/s41577-018-0083-6.
4. D. Nemazee, Mechanisms of central tolerance for B cells. *Nature Reviews Immunology* (2017) https:/doi.org/10.1038/nri.2017.19.
5. Y. Yarkoni, A. Getahun, J. C. Cambier, Molecular underpinning of B-cell anergy. *Immunological Reviews* (2010) https:/doi.org/10.1111/j.1600-065X.2010.00936.x.
6. M. A. ElTanbouly, R. J. Noelle, Rethinking peripheral T cell tolerance: checkpoints across a T cell's journey. *Nature Reviews Immunology* (2021) https:/doi.org/10.1038/s41577-020-00454-2.
7. Z. Yang, *et al.*, Autoimmunity and Traumatic Brain Injury. *Current Physical Medicine and Rehabilitation Reports* (2017) https:/doi.org/10.1007/s40141-017-0146-9.
8. V. M. Stoecklein, A. Osuka, J. A. Lederer, Trauma equals danger--damage control by the immune system. *Journal of Leukocyte Biology* (2012) https:/doi.org/10.1189/jlb.0212072.
9. M. Rojas, *et al.*, Molecular mimicry and autoimmunity. *Journal of Autoimmunity* (2018) https:/doi.org/10.1016/j.jaut.2018.10.012.
10. S. Becart, *et al.*, The role of posttranslational modifications in generating neo-epitopes that bind to rheumatoid arthritis-associated HLA-DR alleles and promote autoimmune T cell responses. *PLoS ONE* (2021) https:/doi.org/10.1371/journal.pone.0245541.
11. C. L. Bennett, *et al.*, The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nature Genetics* (2001) https:/doi.org/10.1038/83713.
12. J. Aaltonen, *et al.*, An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. *Nature Genetics* (1997) https:/doi.org/10.1038/ng1297-399.
13. E. Grip, K. Carlsson, Kostnader för reumatoid artrit i sverige år 2014 (2016).
14. M. Cross, *et al.*, The global burden of rheumatoid arthritis: Estimates from the Global Burden of Disease 2010 study. *Annals of the Rheumatic Diseases* (2014) https:/doi.org/10.1136/annrheumdis-2013-204627.
15. S. T. Ngo, F. J. Steyn, P. A. McCombe, Gender differences in autoimmune disease. *Frontiers in Neuroendocrinology* (2014) https:/doi.org/10.1016/j.yfrne.2014.04.004.
16. H. G. Fassbender, M. Simmling-Annefeld, The potential aggressiveness of synovial tissue in rheumatoid arthritis. *The Journal of Pathology* (1983) https:/doi.org/10.1002/path.1711390314.
17. A. E. Koch, *et al.*, Vascular endothelial growth factor. A cytokine modulating endothelial function in rheumatoid arthritis. *Journal of immunology (Baltimore, Md. : 1950)* (1994).
18. M. Rooney, *et al.*, Analysis of the histologic variation of synovitis in rheumatoid arthritis. *Arthritis & Rheumatism* (1988) https:/doi.org/10.1002/art.1780310803.
19. E. Waaler, On the occurrence of a factor in human serum activating the specific agglutination of sheep blood corpuscles. *APMIS* (2007) https:/doi.org/10.1111/j.1600-0463.2007.apm_682a.x.
20. E. Ossipova, *et al.*, Affinity purified anti-citrullinated protein/peptide antibodies target antigens expressed in the rheumatoid joint. *Arthritis Research and Therapy* **16** (2014).

21. C. Ge, R. Holmdahl, The structure, specificity and function of anti-citrullinated protein antibodies. *Nature Reviews Rheumatology* (2019) https:/doi.org/10.1038/s41584-019-0244-4.
22. K. Nishimura, *et al.*, Meta-analysis: Diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. *Annals of Internal Medicine* (2007) https:/doi.org/10.7326/0003-4819-146-11-200706050-00008.
23. K. Forslind, M. Ahlmén, K. Eberhardt, I. Hafström, B. Svensson, Prediction of radiological outcome in early rheumatoid arthritis in clinical practice: Role of antibodies to citrullinated peptides (anti-CCP). *Annals of the Rheumatic Diseases* (2004) https:/doi.org/10.1136/ard.2003.014233.
24. L. Padyukov, *et al.*, A genome-wide association study suggests contrasting associations in ACPA-positive versus ACPA-negative rheumatoid arthritis. *Annals of the Rheumatic Diseases* (2011) https:/doi.org/10.1136/ard.2009.126821.
25. M. M. J. Nielen, *et al.*, Specific Autoantibodies Precede the Symptoms of Rheumatoid Arthritis: A Study of Serial Measurements in Blood Donors. *Arthritis and Rheumatism* (2004) https:/doi.org/10.1002/art.20018.
26. P. P. Tak, *et al.*, Analysis of the synovial cell infiltrate in early rheumatoid synovial tissue in relation to local disease activity. *Arthritis and Rheumatism* (1997) https:/doi.org/10.1002/art.1780400206.
27. P. P. Tak, *et al.*, Analysis of the synovial cell infiltrate in early rheumatoid synovial tissue in relation to local disease activity. *Arthritis and Rheumatism* (1997) https:/doi.org/10.1002/art.1780400206.
28. Y. Okada, S. Eyre, A. Suzuki, Y. Kochi, K. Yamamoto, Genetics of rheumatoid arthritis: 2018 status. *Annals of the Rheumatic Diseases* (2019) https:/doi.org/10.1136/annrheumdis-2018-213678.
29. D. Baeten, *et al.*, Comparative study of the synovial histology in rheumatoid arthritis, spondyloarthropathy, and osteoarthritis: Influence of disease duration and activity. *Annals of the Rheumatic Diseases* (2000) https:/doi.org/10.1136/ard.59.12.945.
30. J. A. van Boxel, S. A. Paget, Predominantly T-Cell Infiltrate in Rheumatoid Synovial Membranes. *New England Journal of Medicine* (1975) https:/doi.org/10.1056/nejm197509112931101.
31. A. E. Schröder, A. Greiner, C. Seyfert, C. Berek, Differentiation of B cells in the nonlymphoid tissue of the synovial membrane of patients with rheumatoid arthritis. *Proceedings of the National Academy of Sciences of the United States of America* (1996) https:/doi.org/10.1073/pnas.93.1.221.
32. S. Takemura, *et al.*, Lymphoid Neogenesis in Rheumatoid Synovitis. *The Journal of Immunology* (2001) https:/doi.org/10.4049/jimmunol.167.2.1072.
33. H. U. Scherer, T. W. J. Huizinga, G. Krönke, G. Schett, R. E. M. Toes, The B cell response to citrullinated antigens in the development of rheumatoid arthritis. *Nature Reviews Rheumatology* (2018) https:/doi.org/10.1038/nrrheum.2018.10.
34. J. M. Kremer, *et al.*, Treatment of Rheumatoid Arthritis by Selective Inhibition of T-Cell Activation with Fusion Protein CTLA4Ig. *New England Journal of Medicine* (2003) https:/doi.org/10.1056/nejmoa035075.
35. J. C. W. Edwards, *et al.*, Efficacy of B-Cell–Targeted Therapy with Rituximab in Patients with Rheumatoid Arthritis. *New England Journal of Medicine* (2004) https:/doi.org/10.1056/nejmoa032534.
36. R. M. Thurlings, *et al.*, Synovial tissue response to rituximab: Mechanism of action and identification of biomarkers of response. *Annals of the Rheumatic Diseases* **67** (2008).
37. M. E. Duddy, A. Alter, A. Bar-Or, Distinct Profiles of Human B Cell Effector Cytokines: A Role in Immune Regulation? *The Journal of Immunology* **172** (2004).

38. S. L. Constant, B lymphocytes as antigen-presenting cells for CD4+ T cell priming in vivo. *Journal of immunology (Baltimore, Md. : 1950)* (1999).
39. S. Takemura, P. A. Klimiuk, A. Braun, J. J. Goronzy, C. M. Weyand, T Cell Activation in Rheumatoid Synovium Is B Cell Dependent. *The Journal of Immunology* **167** (2001).
40. R. Thomas, M. McIlraith, L. S. Davis, P. E. Lipsky, Rheumatoid synovium is enriched in CD45RBdim mature memory T cells that are potent helpers for B cell differentiation. *Arthritis & Rheumatism* (1992) https:/doi.org/10.1002/art.1780351209.
41. P. L. Klarenbeek, *et al.*, Inflamed target tissue provides a specific niche for highly expanded T-cell clones in early human autoimmune disease. *Annals of the Rheumatic Diseases* (2012) https:/doi.org/10.1136/annrheumdis-2011-200612.
42. E. A. James, *et al.*, Citrulline-specific Th1 cells are increased in rheumatoid arthritis and their frequency is influenced by disease duration and therapy. *Arthritis and Rheumatology* (2014) https:/doi.org/10.1002/art.38637.
43. N. Matthews, P. Emery, D. Pilling, A. Akbar, M. Salmon, Subpopulations of primed t helper cells in rheumatoid arthritis. *Arthritis & Rheumatism* (1993) https:/doi.org/10.1002/art.1780360505.
44. K. J. Warrington, S. Takemura, J. J. Goronzy, C. M. Weyand, CD4+,CD28- T cells in rheumatoid arthritis patients combine features of the innate and adaptive immune systems. *Arthritis and Rheumatism* (2001) https:/doi.org/10.1002/1529-0131(200101)44:1<13::AID-ANR3>3.0.CO;2-6.
45. A. P. Cope, H. Schulze-Koops, M. Aringer, The central role of T cells in rheumatoid arthritis. *Clinical and Experimental Rheumatology* (2007).
46. K. Raza, *et al.*, Synovial fluid leukocyte apoptosis is inhibited in patients with very early rheumatoid arthritis. *Arthritis Research and Therapy* **8** (2006).
47. R. J. E. M. Dolhain, A. N. van der Heiden, N. T. ter Haar, F. C. Breedveld, A. M. M. Miltenburg, Shift toward T lymphocytes with a T helper 1 cytokine-secretion profile in the joints of patients with rheumatoid arthritis. *Arthritis and Rheumatism* **39** (1996).
48. S. Yang, C. Smith, J. M. Prahl, X. Luo, H. F. DeLuca, Vitamin D Deficiency Suppresses Cell-Mediated Immunity in Vivo. *Archives of Biochemistry and Biophysics* (1993) https:/doi.org/10.1006/abbi.1993.1260.
49. M. Chabaud, *et al.*, Human interleukin-17: A T cell-derived proinflammatory cytokine produced by the rheumatoid synovium. *Arthritis and Rheumatism* (1999) https:/doi.org/10.1002/1529-0131(199905)42:5<963::AID-ANR15>3.0.CO;2-E.
50. H. Yamada, *et al.*, Th1 but not Th17 cells predominate in the joints of patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases* **67** (2008).
51. P. A. Klimiuk, H. Yang, J. J. Goronzy, C. M. Weyand, Production of cytokines and metalloproteinases in rheumatoid synovitis is T cell dependent. *Clinical Immunology* (1999) https:/doi.org/10.1006/clim.1998.4618.
52. S. Kotake, *et al.*, IL-17 in synovial fluids from patients with rheumatoid arthritis is a potent stimulator of osteoclastogenesis. *Journal of Clinical Investigation* (1999) https:/doi.org/10.1172/JCI5703.
53. M. Takeshita, *et al.*, Multi-dimensional analysis identified rheumatoid arthritis-driving pathway in human T cell. *Annals of the Rheumatic Diseases* **78** (2019).
54. J. Tuncel, *et al.*, Self-reactive T cells induce and perpetuate chronic relapsing arthritis. *Arthritis Research and Therapy* (2020) https:/doi.org/10.1186/s13075-020-2104-7.
55. I. Cecchi, *et al.*, Neutrophils: Novel key players in Rheumatoid Arthritis. Current and future therapeutic targets. *Autoimmunity Reviews* **17** (2018).
56. I. A. Udalova, A. Mantovani, M. Feldmann, Macrophage heterogeneity in the context of rheumatoid arthritis. *Nature Reviews Rheumatology* **12** (2016).

57. S. Culemann, *et al.*, Locally renewing resident synovial macrophages provide a protective barrier for the joint. *Nature* (2019) https:/doi.org/10.1038/s41586-019-1471-1.
58. K. Shin, *et al.*, Mast Cells Contribute to Autoimmune Inflammatory Arthritis via Their Tryptase/Heparin Complexes. *The Journal of Immunology* **182** (2009).
59. F. Rivellese, F. W. Rossi, M. R. Galdiero, C. Pitzalis, A. de Paulis, Mast cells in early rheumatoid arthritis. *International Journal of Molecular Sciences* **20** (2019).
60. J. Suurmond, *et al.*, Communication between human mast cells and CD4+ T cells through antigen-dependent interactions. *European Journal of Immunology* **43** (2013).
61. A. J. Hueber, *et al.*, Cutting Edge: Mast Cells Express IL-17A in Rheumatoid Arthritis Synovium. *The Journal of Immunology* **184** (2010).
62. B. Bartok, G. S. Firestein, Fibroblast-like synoviocytes: Key effector cells in rheumatoid arthritis. *Immunological Reviews* **233** (2010).
63. C. C. Tseng, *et al.*, Dual role of chondrocytes in rheumatoid arthritis: The chicken and the egg. *International Journal of Molecular Sciences* **21** (2020).
64. N. M. T. Roodenrijs, *et al.*, Characteristics of difficult-to-treat rheumatoid arthritis: Results of an international survey. *Annals of the Rheumatic Diseases* (2018) https:/doi.org/10.1136/annrheumdis-2018-213687.
65. T. Bongartz, *et al.*, Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: Systematic review and meta-analysis of rare harmful effects in randomized controlled trials. *Journal of the American Medical Association* (2006) https:/doi.org/10.1001/jama.295.19.2275.
66. S. Viatte, A. Barton, Genetics of rheumatoid arthritis susceptibility, severity, and treatment response. *Seminars in Immunopathology* (2017) https:/doi.org/10.1007/s00281-017-0630-4.
67. L. Klareskog, V. Malmström, K. Lundberg, L. Padyukov, L. Alfredsson, Smoking, citrullination and genetic variability in the immunopathogenesis of rheumatoid arthritis. *Seminars in Immunology* **23**, 92–98 (2011).
68. N. Wegner, *et al.*, Peptidylarginine deiminase from Porphyromonas gingivalis citrullinates human fibrinogen and α-enolase: Implications for autoimmunity in rheumatoid arthritis. *Arthritis & Rheumatism* **62**, 2662–2672 (2010).
69. M. M. Newkirk, K. N. W. Duffy, J. Leclerc, N. Lambert, J. B. Shiroky, Detection of cytomegalovirus, epstein-barr virus and herpes virus-6 in patients with rheumatoid arthritis with or without sjögren's syndrome. *Rheumatology* (1994) https:/doi.org/10.1093/rheumatology/33.4.317.
70. S. R. Harrison, D. Li, L. E. Jeffery, K. Raza, M. Hewison, Vitamin D, Autoimmune Disease and Rheumatoid Arthritis. *Calcified Tissue International* (2019) https:/doi.org/10.1007/s00223-019-00577-2.
71. C. Orellana, *et al.*, Oral contraceptives, breastfeeding and the risk of developing rheumatoid arthritis: Results from the Swedish EIRA study. *Annals of the Rheumatic Diseases* (2017) https:/doi.org/10.1136/annrheumdis-2017-211620.
72. B. Sverdrup, *et al.*, Association between occupational exposure to mineral oil and rheumatoid arthritis: results from the Swedish EIRA case-control study. *Arthritis research & therapy* **7**, R1296-303 (2005).
73. N. Balandraud, J. Roudier, Epstein-Barr virus and rheumatoid arthritis. *Joint Bone Spine* (2018) https:/doi.org/10.1016/j.jbspin.2017.04.011.
74. J. Roudier, J. Petersen, G. H. Rhodes, J. Luka, D. A. Carson, Susceptibility to rheumatoid arthritis maps to a T-cell epitope shared by the HLA-Dw4 DR β-1 chain and the Epstein-Barr virus glycoprotein gp110. *Proceedings of the National Academy of Sciences of the United States of America* (1989) https:/doi.org/10.1073/pnas.86.13.5104.

75. J. Roudier, J. Petersen, G. H. Rhodes, J. Luka, D. A. Carson, Susceptibility to rheumatoid arthritis maps to a T-cell epitope shared by the HLA-Dw4 DR β-1 chain and the Epstein-Barr virus glycoprotein gp110. *Proceedings of the National Academy of Sciences of the United States of America* (1989) https:/doi.org/10.1073/pnas.86.13.5104.
76. S. W. Scally, *et al.*, A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *Journal of Experimental Medicine* (2013) https:/doi.org/10.1084/jem.20131241.
77. Y. Kochi, *et al.*, PADI4 polymorphism predisposes male smokers to rheumatoid arthritis. *Annals of the Rheumatic Diseases* (2011) https:/doi.org/10.1136/ard.2010.130526.
78. D. Makrygiannakis, *et al.*, Smoking increases peptidylarginine deiminase 2 enzyme expression in human lungs and increases citrullination in BAL cells. *Annals of the Rheumatic Diseases* (2008) https:/doi.org/10.1136/ard.2007.075192.
79. E. B. Lugli, *et al.*, Expression of citrulline and homocitrulline residues in the lungs of non-smokers and smokers: Implications for autoimmunity in rheumatoid arthritis. *Arthritis Research and Therapy* (2015) https:/doi.org/10.1186/s13075-015-0520-x.
80. W. Nesse, *et al.*, The periodontium of periodontitis patients contains citrullinated proteins which may play a role in ACPA (anti-citrullinated protein antibody) formation. *Journal of Clinical Periodontology* (2012) https:/doi.org/10.1111/j.1600-051X.2012.01885.x.
81. K. Eriksson, *et al.*, Seropositivity combined with smoking is associated with increased prevalence of periodontitis in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases* (2018) https:/doi.org/10.1136/annrheumdis-2017-212091.
82. I. Kostoglou-Athanassiou, P. Athanassiou, A. Lyraki, I. Raftakis, C. Antoniadis, Vitamin D and rheumatoid arthritis. *Therapeutic Advances in Endocrinology and Metabolism* (2012) https:/doi.org/10.1177/2042018812471070.
83. L. A. Merlino, *et al.*, Vitamin D Intake Is Inversely Associated With Rheumatoid Arthritis: Results From the Iowa Women's Health Study. *Arthritis and Rheumatism* **50**, 72–77 (2004).
84. G. G. Song, S. C. Bae, Y. H. Lee, Association between vitamin D intake and the risk of rheumatoid arthritis: A meta-analysis. *Clinical Rheumatology* (2012) https:/doi.org/10.1007/s10067-012-2080-7.
85. M. Pedersen, *et al.*, Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis Research and Therapy* (2006) https:/doi.org/10.1186/ar2022.
86. M. Pedersen, *et al.*, Strong combined gene-environment effects in anti-cyclic citrullinated peptide-positive rheumatoid arthritis: A nationwide case-control study in Denmark. *Arthritis and Rheumatism* (2007) https:/doi.org/10.1002/art.22597.
87. S. T. Ngo, F. J. Steyn, P. A. McCombe, Gender differences in autoimmune disease. *Frontiers in Neuroendocrinology* **35**, 347–369 (2014).
88. L. A. Merlino, J. R. Cerhan, L. A. Criswell, T. R. Mikuls, K. G. Saag, Estrogen and Other Female Reproductive Risk Factors Are Not Strongly Associated With the Development of Rheumatoid Arthritis in Elderly Women. *Seminars in Arthritis and Rheumatism* (2003) https:/doi.org/10.1016/S0049-0172(03)00084-2.
89. G. M. Hall, M. Daniels, E. C. Huskisson, T. D. Spector, A randomised controlled trial of the effect of hormone replacement therapy on disease activity in postmenopausal rheumatoid arthritis. *Annals of the Rheumatic Diseases* (1994) https:/doi.org/10.1136/ard.53.2.112.
90. K. P. Liao, L. Alfredsson, E. W. Karlson, Environmental influences on risk for rheumatoid arthritis. *Current opinion in rheumatology* **21**, 279–83 (2009).

91. P. Stastny, Mixed lymphocyte cultures in rheumatoid arthritis. *Journal of Clinical Investigation* (1976) https:/doi.org/10.1172/JCI108382.
92. P. K. Gregersen, J. Silver, R. J. Winchester, The Shared Epitope Hypothesis. *Arthritis and Rheumatism* (1987).
93. Y. Okada, S. Eyre, A. Suzuki, Y. Kochi, K. Yamamoto, Genetics of rheumatoid arthritis: 2018 status. *Annals of the Rheumatic Diseases* (2019) https:/doi.org/10.1136/annrheumdis-2018-213678.
94. K. McAllister, S. Eyre, G. Orozco, Genetics of rheumatoid arthritis: GWAS and beyond. *Open Access Rheumatology: Research and Reviews* (2011) https:/doi.org/10.2147/OARRR.S14725.
95. Y. Okada, S. Eyre, A. Suzuki, Y. Kochi, K. Yamamoto, Genetics of rheumatoid arthritis: 2018 status. *Annals of the Rheumatic Diseases* (2019) https:/doi.org/10.1136/annrheumdis-2018-213678.
96. L. Jiang, *et al.*, Novel risk loci for rheumatoid arthritis in han chinese and congruence with risk variants in europeans. *Arthritis and Rheumatology* (2014) https:/doi.org/10.1002/art.38353.
97. S. Viatte, A. Barton, Genetics of rheumatoid arthritis susceptibility, severity, and treatment response. *Seminars in Immunopathology* (2017) https:/doi.org/10.1007/s00281-017-0630-4.
98. D. Van Der Woude, *et al.*, Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. *Arthritis and Rheumatism* (2009) https:/doi.org/10.1002/art.24385.
99. F. Kurreeman, *et al.*, Genetic basis of autoantibody positive and negative rheumatoid arthritis risk in a multi-ethnic cohort derived from electronic health records. *American Journal of Human Genetics* (2011) https:/doi.org/10.1016/j.ajhg.2010.12.007.
100. P. L. Auer, G. Lettre, Rare variant association studies: Considerations, challenges and opportunities. *Genome Medicine* (2015) https:/doi.org/10.1186/s13073-015-0138-2.
101. T. A. Manolio, *et al.*, Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
102. L. M. Olsson, *et al.*, A single nucleotide polymorphism in the NCF1 gene leading to reduced oxidative burst is associated with systemic lupus erythematosus. *Annals of the Rheumatic Diseases* (2017) https:/doi.org/10.1136/annrheumdis-2017-211287.
103. J. Yang, *et al.*, Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* (2010) https:/doi.org/10.1038/ng.608.
104. T. F. C. Mackay, R. R. H. Anholt, Of flies and man: Drosophila as a model for human complex traits. *Annual Review of Genomics and Human Genetics* (2006) https:/doi.org/10.1146/annurev.genom.7.080505.115758.
105. L. Bomba, K. Walter, N. Soranzo, The impact of rare and low-frequency genetic variants in common disease. *Genome Biology* **18**, 77 (2017).
106. O. A. Panagiotou, *et al.*, What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. *International Journal of Epidemiology* (2012) https:/doi.org/10.1093/ije/dyr178.
107. P. C. Sham, S. M. Purcell, Statistical power and significance testing in large-scale genetic studies. *Nature Reviews Genetics* **15**, 335–346 (2014).
108. B. Elsworth, *et al.*, The MRC IEU OpenGWAS data infrastructure. *bioRxiv* (2020) https:/doi.org/10.1101/2020.08.10.244293.
109. S. L. Edwards, J. Beesley, J. D. French, A. M. Dunning, Beyond GWASs: illuminating the dark road from association to function. *American journal of human genetics* **93**, 779–97 (2013).
110. M. T. Maurano, *et al.*, Systematic localization of common disease-associated variation in regulatory DNA. *Science (New York, N.Y.)* **337**, 1190–5 (2012).

111. K. Musunuru, *et al.*, From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature* **466**, 714–9 (2010).
112. O. Harismendy, *et al.*, 9p21 DNA variants associated with coronary artery disease impair interferon-γ signalling response. *Nature* **470**, 264–8 (2011).
113. A. M. Walsh, *et al.*, Integrative genomic deconvolution of rheumatoid arthritis GWAS loci into gene and cell type associations. *Genome biology* **17**, 79 (2016).
114. K. Paigen, A miracle enough: The power of mice. *Nature Medicine* (1995) https:/doi.org/10.1038/nm0395-215.
115. R. H. Waterston, *et al.*, Initial sequencing and comparative analysis of the mouse genome. *Nature* (2002) https:/doi.org/10.1038/nature01262.
116. J. Mestas, C. C. W. Hughes, Of Mice and Not Men: Differences between Mouse and Human Immunology. *The Journal of Immunology* (2004) https:/doi.org/10.4049/jimmunol.172.5.2731.
117. K. Takao, T. Miyakawa, Genomic responses in mouse models greatly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences of the United States of America* (2015) https:/doi.org/10.1073/pnas.1401965111.
118. Junhee Seok, *et al.*, Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences of the United States of America* (2013) https:/doi.org/10.1073/pnas.1222878110.
119. D. T. Burke, *et al.*, Dissection of complex adult traits in a mouse synthetic population. *Genome research* **22**, 1549–57 (2012).
120. J. Flint, E. Eskin, Genome-wide association studies in mice. *Nature reviews. Genetics* **13**, 807–17 (2012).
121. A.-K. B. Lindqvist, *et al.*, Mouse models for rheumatoid arthritis. *Trends in Genetics* **18**, S7–S13 (2002).
122. S. Lu, S. Carlsen, A. S. Hansson, R. Holmdahl, Immunization of rats with homologous type XI collagen leads to chronic and relapsing arthritis with different genetics and joint pathology than arthritis induced with homologous type II collagen. *Journal of Autoimmunity* (2002) https:/doi.org/10.1006/jaut.2001.0581.
123. S. Carlsén, A. S. Hansson, H. Olsson, D. Heinegård, R. Holmdahl, Cartilage oligomeric matrix protein (COMP)-induced arthritis in rats. *Clinical and Experimental Immunology* (1998) https:/doi.org/10.1046/j.1365-2249.1998.00739.x.
124. K. Iwanami, *et al.*, Arthritogenic T cell epitope in glucose-6-phosphate isomerase-induced arthritis. *Arthritis Research & Therapy* **10**, R130 (2008).
125. C. Vingsbo, *et al.*, Pristane-induced arthritis in rats: A new model for rheumatoid arthritis with a chronic disease course influenced by both major histocompatibility complex and non-major histocompatibility complex genes. *American Journal of Pathology* (1996).
126. P. Li, E. M. Schwarz, The TNF-α transgenic mouse model of inflammatory arthritis. *Springer Seminars in Immunopathology* (2003) https:/doi.org/10.1007/s00281-003-0125-3.
127. N. Sakaguchi, *et al.*, Altered thymic T-cell selection due to a mutation of the ZAP-70 gene causes autoimmune arthritis in mice. *Nature* **426**, 454–460 (2003).
128. D. D. Brand, K. A. Latham, E. F. Rosloniec, Collagen-induced arthritis. *Nature Protocols* (2007) https:/doi.org/10.1038/nprot.2007.173.
129. T. J. Goldschmidt, R. Holmdahl, Anti-T cell receptor antibody treatment of rats with established autologous collagen-induced arthritis: suppression of arthritis without reduction of anti-type II collagen autoantibody levels. *European Journal of Immunology* (1991) https:/doi.org/10.1002/eji.1830210536.
130. K. M. Kadowaki, H. Matsuno, H. Tsuji, I. Tunru, CD4+ T cells from collagen-induced arthritic mice are essential to transfer arthritis into severe combined immunodeficient

mice. *Clinical and Experimental Immunology* (1994) https:/doi.org/10.1111/j.1365-2249.1994.tb06070.x.
131. A. Corthay, A. Johansson, M. Vestberg, R. Holmdahl, Collagen-induced arthritis development requires alpha beta T cells but not gamma delta T cells: studies with T cell-deficient (TCR mutant) mice. *International immunology* **11**, 1065–73 (1999).
132. K. S. Nandakumar, J. Bäcklund, M. Vestberg, R. Holmdahl, Collagen type II (CII)-specific antibodies induce arthritis in the absence of T or B cells but the arthritis progression is enhanced by CII-reactive T cells. *Arthritis research & therapy* **6**, R544–R550 (2004).
133. S. H. Park, *et al.*, BATF regulates collagen-induced arthritis by regulating T helper cell differentiation. *Arthritis Research and Therapy* (2018) https:/doi.org/10.1186/s13075-018-1658-0.
134. P. Caplazi, *et al.*, Mouse Models of Rheumatoid Arthritis. *Veterinary Pathology* **52**, 819–826 (2015).
135. J. S. Courtenay, M. J. Dallman, A. D. Dayan, A. Martin, B. Mosedale, Immunisation against heterologous type II collagen induces arthritis in mice [13]. *Nature* (1980) https:/doi.org/10.1038/283666a0.
136. S. Nakae, A. Nambu, K. Sudo, Y. Iwakura, Suppression of Immune Induction of Collagen-Induced Arthritis in IL-17-Deficient Mice. *The Journal of Immunology* (2003) https:/doi.org/10.4049/jimmunol.171.11.6173.
137. M. -C Boissier, *et al.*, Biphasic effect of interferon-γ in murine collagen-induced arthritis. *European Journal of Immunology* (1995) https:/doi.org/10.1002/eji.1830250508.
138. T. Alonzi, *et al.*, Interleukin 6 is required for the development of collagen-induced arthritis. *Journal of Experimental Medicine* (1998) https:/doi.org/10.1084/jem.187.4.461.
139. D. M. Butler, *et al.*, DBA/1 mice expressing the human TNF-alpha transgene develop a severe, erosive arthritis: characterization of the cytokine cascade and cellular composition. *Journal of immunology (Baltimore, Md. : 1950)* (1997).
140. W. B. van den Berg, L. A. B. Joosten, M. Helsen, F. A. J. van de Loo, Amelioration of established murine collagen-induced arthritis with anti- IL-1 treatment. *Clinical and Experimental Immunology* (1994) https:/doi.org/10.1111/j.1365-2249.1994.tb06517.x.
141. H. O. Mcdevitt, A. Chinitz, Genetic control of the antibody response: Relationship between immune response and histocompatibility (H-2) type. *Science* (1969) https:/doi.org/10.1126/science.163.3872.1207.
142. E. Ahlqvist, R. Bockermann, R. Holmdahl, Fragmentation of two quantitative trait loci controlling collagen-induced arthritis reveals a new set of interacting subloci. *Journal of immunology (Baltimore, Md. : 1950)* **178**, 3084–90 (2007).
143. M. Johannesson, *et al.*, Gene expression profiling of arthritis using a QTL chip reveals a complex gene regulation of the Cia5 region in mice. *Genes and immunity* **6**, 575–83 (2005).
144. A. Darvasi, Experimental strategies for the genetic dissection of complex traits in animal models. *Nature Genetics* (1998) https:/doi.org/10.1038/ng0198-19.
145. G. A. Churchill, *et al.*, The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nature Genetics* (2004) https:/doi.org/10.1038/ng1104-1133.
146. J. Flint, W. Valdar, S. Shifman, R. Mott, Strategies for mapping and cloning quantitative trait genes in rodents. *Nature Reviews Genetics* **6**, 271–286 (2005).
147. S. Shifman, A. Darvasi, Mouse inbred strain sequence information and yin-yang crosses for quantitative trait locus fine mapping. *Genetics* (2005) https:/doi.org/10.1534/genetics.104.032474.

148. W. Valdar, *et al.*, Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nature Genetics* (2006) https:/doi.org/10.1038/ng1840.
149. G. A. Churchill, D. M. Gatti, S. C. Munger, K. L. Svenson, The Diversity Outbred mouse population. *Mammalian Genome* (2012) https:/doi.org/10.1007/s00335-012-9414-2.
150. G. Manenti, *et al.*, Mouse genome-wide association mapping needs linkage analysis to avoid false-positive loci. *PLoS Genetics* (2009) https:/doi.org/10.1371/journal.pgen.1000331.
151. D. Vaartjes, *et al.*, Genetic dissection of a major haplotype associated with arthritis reveal FcγR2b and FcγR3 to act additively. *European Journal of Immunology* (2020) https:/doi.org/10.1002/eji.202048605.
152. N. R. Drinkwater, *et al.*, The Long Path from QTL to Gene. *PLoS Genetics* **8**, e1002975 (2012).
153. S. T. Qureshi, *et al.*, Endotoxin-tolerant mice have mutations in toll-like receptor 4 (Tlr4). *Journal of Experimental Medicine* (1999) https:/doi.org/10.1084/jem.189.4.615.
154. A. A. Perelygin, *et al.*, Positional cloning of the murine flavivirus resistance gene. *Proceedings of the National Academy of Sciences of the United States of America* (2002) https:/doi.org/10.1073/pnas.142287799.
155. P. Olofsson, *et al.*, Positional identification of Ncf1 as a gene that regulates arthritis severity in rats. *Nature Genetics* (2003) https:/doi.org/10.1038/ng1058.
156. J. R. Shorter, *et al.*, Quantitative trait mapping in Diversity Outbred mice identifies two genomic regions associated with heart size. *Mammalian Genome* (2018) https:/doi.org/10.1007/s00335-017-9730-7.
157. U. C. Rogner, P. Avner, Congenic mice: cutting tools for complex immune disorders. *Nature Reviews Immunology* **3**, 243–252 (2003).
158. T. L. Turner, Fine-mapping natural alleles: Quantitative complementation to the rescue. *Molecular Ecology* (2014) https:/doi.org/10.1111/mec.12719.
159. K. E. Noll, M. T. Ferris, M. T. Heise, The Collaborative Cross: A Systems Genetics Resource for Studying Host-Pathogen Interactions. *Cell Host and Microbe* (2019) https:/doi.org/10.1016/j.chom.2019.03.009.
160. M. Johannesson, *et al.*, Identification of epistasis through a partial advanced intercross reveals three arthritis loci within the Cia5 QTL in mice. *Genes and Immunity* **6**, 175–185 (2005).
161. U. Norin, *et al.*, Endophilin A2 deficiency protects rodents from autoimmune arthritis by modulating T cell activation. *Nature Communications* (2021) https:/doi.org/10.1038/s41467-020-20586-2.
162. C. Vingsbo, *et al.*, Pristane-induced arthritis in rats: A new model for rheumatoid arthritis with a chronic disease course influenced by both major histocompatibility complex and non-major histocompatibility complex genes. *American Journal of Pathology* (1996).
163. A. Pizzolla, F. Laulund, K. Wing, R. Holmdahl, A new model of arthritis induced by a glucose-6-phosphate isomerase peptide: immunological requirements and peptide characterisation. *Annals of the Rheumatic Diseases* **71**, A83.2-A83 (2012).
164. M. Schaller, D. R. Burton, H. J. Ditzel, Autoantibodies to GPI in rheumatoid arthritis: linkage between an animal model and human disease. *Nature immunology* **2**, 746–53 (2001).
165. K. S. Nandakumar, R. Holmdahl, Collagen antibody induced arthritis. *Methods in molecular medicine* (2007) https:/doi.org/10.1007/978-1-59745-402-5_16.
166. Y. Luo, M. E. Dorf, Delayed-type hypersensitivity. *Current protocols in immunology / edited by John E. Coligan ... [et al.]* **Chapter 4**, Unit 4.5 (2001).

167. S. Glatigny, E. Bettelli, Experimental Autoimmune Encephalomyelitis (EAE) as Animal Models of Multiple Sclerosis (MS). *Cold Spring Harbor perspectives in medicine* (2018) https:/doi.org/10.1101/cshperspect.a028977.
168. B. Chassaing, J. D. Aitken, M. Malleshappa, M. Vijay-Kumar, Dextran sulfate sodium (DSS)-induced colitis in mice. *Current protocols in immunology / edited by John E. Coligan ... [et al.]* **104**, Unit 15.25. (2014).
169. M. Green, Cod liver oil and tuberculosis. *BMJ (Online)* (2011) https:/doi.org/10.1136/bmj.d7505.
170. P. T. Liu, *et al.*, Toll-like receptor triggering of a vitamin D-mediated human antimicrobial response. *Science* (2006) https:/doi.org/10.1126/science.1123933.
171. T. Skaaby, L. L. N. Husemoen, B. H. Thuesen, A. Linneberg, Prospective population-based study of the association between vitamin D status and incidence of autoimmune disease. *Endocrine* (2015) https:/doi.org/10.1007/s12020-015-0547-4.
172. Y. H. Lee, S. C. Bae, S. J. Choi, J. D. Ji, G. G. Song, Associations between vitamin D receptor polymorphisms and susceptibility to rheumatoid arthritis and systemic lupus erythematosus: A meta-analysis. *Molecular Biology Reports* (2011) https:/doi.org/10.1007/s11033-010-0477-4.
173. G. B. Ferreira, *et al.*, 1,25-Dihydroxyvitamin D 3 Promotes Tolerogenic Dendritic Cells with Functional Migratory Properties in NOD Mice . *The Journal of Immunology* (2014) https:/doi.org/10.4049/jimmunol.1302350.
174. W. F. C. Rigby, T. Stacy, M. W. Fanger, Inhibition of T lymphocyte mitogenesis by 1,25-dihydroxyvitamin D3 (calcitriol). *Journal of Clinical Investigation* (1984) https:/doi.org/10.1172/JCI111557.
175. L. E. Jeffery, *et al.*, 1,25-Dihydroxyvitamin D3 and IL-2 Combine to Inhibit T Cell Production of Inflammatory Cytokines and Promote Development of Regulatory T Cells Expressing CTLA-4 and FoxP3. *The Journal of Immunology* (2009) https:/doi.org/10.4049/jimmunol.0803217.
176. D. Bruce, S. Yu, J. H. Ooi, M. T. Cantorna, Converging pathways lead to overproduction of IL-17 in the absence of vitamin D signaling. *International Immunology* (2011) https:/doi.org/10.1093/intimm/dxr045.
177. K. E. Hansen, C. M. Bartels, R. E. Gangnon, A. N. Jones, J. Gogineni, An evaluation of high-dose vitamin D for rheumatoid arthritis. *Journal of clinical rheumatology : practical reports on rheumatic & musculoskeletal diseases* **20**, 112–4 (2014).
178. M. Salesi, Z. Farajzadegan, Efficacy of Vitamin D in patients with active rheumatoid arthritis receiving methotrexate therapy. *Rheumatology International* **32**, 2129–2133 (2012).
179. W. Dankers, E. M. Colin, J. P. van Hamburg, E. Lubberts, Vitamin D in Autoimmunity: Molecular Mechanisms and Therapeutic Potential. *Frontiers in immunology* **7**, 697 (2016).
180. E. C. Keystone, Abandoned therapies and unpublished trials in rheumatoid arthritis. *Current Opinion in Rheumatology* (2003) https:/doi.org/10.1097/00002281-200305000-00012.
181. K. S. Rosenthal, R. Carambula, D. H. Zimmerman, Why Don't We Have a Vaccine Against Autoimmune Diseases? - A Review. *Journal of clinical & cellular immunology* (2019) https:/doi.org/10.4172/2155-9899.1000574.
182. M. L. Barnett, *et al.*, Treatment of rheumatoid arthritis with oral type II collagen: Results of a multicenter, double-blind, placebo-controlled trial. *Arthritis and Rheumatism* (1998) https:/doi.org/10.1002/1529-0131(199802)41:2<290::AID-ART13>3.0.CO;2-R.
183. M. Cazzola, *et al.*, Oral type II collagen in the treatment of rheumatoid arthritis. A six-month double blind placebo-controlled study. *Clinical and Experimental Rheumatology* (2000).

184. D. Diogo, *et al.*, Rare, low-frequency, and common variants in the protein-coding sequence of biological candidate genes from GWASs contribute to risk of rheumatoid arthritis. *American Journal of Human Genetics* (2013) https:/doi.org/10.1016/j.ajhg.2012.11.012.
185. S. Raychaudhuri, *et al.*, Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. *Nature Genetics* (2009) https:/doi.org/10.1038/ng.479.
186. M. Østensen, L. Fuhrer, R. Mathieu, M. Seitz, P. M. Villiger, A prospective study of pregnant patients with rheumatoid arthritis and ankylosing spondylitis using validated clinical instruments. *Annals of the Rheumatic Diseases* (2004) https:/doi.org/10.1136/ard.2003.016881.
187. P. Dieppe, F. A. Wollheim, H. R. Schumacher, "The ameliorating effect of pregnancy on chronic atrophic (infectious rheumatoid) arthritis, fibrositis, and intermittent hydrarthrosis" in *Classic Papers in Rheumatology*, (2021) https:/doi.org/10.3109/9780203214237-49.
188. N. S. Latman, Relation of menstrual cycle phase to symptoms of rheumatoid arthritis. *The American Journal of Medicine* (1983) https:/doi.org/10.1016/0002-9343(83)90789-1.
189. G. M. Hall, M. Daniels, E. C. Huskisson, T. D. Spector, A randomised controlled trial of the effect of hormone replacement therapy on disease activity in postmenopausal rheumatoid arthritis. *Annals of the Rheumatic Diseases* (1994) https:/doi.org/10.1136/ard.53.2.112.
190. I. F. Cook, Sexual dimorphism of humoral immunity with human vaccines. *Vaccine* (2008) https:/doi.org/10.1016/j.vaccine.2008.04.054.
191. K. L. Phiel, R. A. Henderson, S. J. Adelman, M. M. Elloso, Differential estrogen receptor gene expression in human peripheral blood mononuclear cell populations. *Immunology Letters* (2005) https:/doi.org/10.1016/j.imlet.2004.10.007.
192. W. Gilmore, L. P. Weiner, J. Correale, Effect of estradiol on cytokine secretion by proteolipid protein-specific T cell clones isolated from multiple sclerosis patients and normal control subjects. *Journal of immunology (Baltimore, Md. : 1950)* (1997).
193. A. Maret, *et al.*, Estradiol enhances primary antigen-specific CD4 T cell responses and Th1 development in vivo. Essential role of estrogen receptor α expression in hematopoietic cells. *European Journal of Immunology* (2003) https:/doi.org/10.1002/immu.200310027.
194. A. Andersson, *et al.*, Estrogen regulates T helper 17 phenotype and localization in experimental autoimmune arthritis. *Arthritis research & therapy* **17**, 32 (2015).
195. P. L. Podolin, *et al.*, Localization of two insulin-dependent diabetes (Idd) genes to the Idd10 region on mouse Chromosome 3. *Mammalian Genome* (1998) https:/doi.org/10.1007/s003359900749.
196. M. Bagot, *et al.*, CD101 is expressed by skin dendritic cells. Role in T-lymphocyte activation. *Tissue antigens* **50**, 439–48 (1997).
197. G. Fernandez Lahore, *et al.*, Vitamin D3 receptor polymorphisms regulate T cells and T cell-dependent inflammatory diseases. *Proceedings of the National Academy of Sciences* (2020) https:/doi.org/10.1073/pnas.2001966117.
198. G. Fernandez Lahore, *et al.*, Polymorphic estrogen receptor binding site causes CD2-dependent sex bias in the susceptibility to autoimmune diseases. *Nature Portfolio* (2021) https:/doi.org/10.21203/rs.3.rs-337166/v1.
199. U. Norin, *et al.*, Endophilin A2 deficiency protects rodents from autoimmune arthritis by modulating T cell activation. *Nature Communications* (2021) https:/doi.org/10.1038/s41467-020-20586-2.